



Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinska,M.I., Skalska,U., Snailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.,  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 3326)  
Strausberg,R.  
Direct Submission  
Submitted (29-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Sep 16, 2003 this sequence version replaced gi:26251822.  
Contact: MGC help desk  
Email: [cgabs@remail.nih.gov](mailto:cgabs@remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toehiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nees, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 70 Row: p Column: 2.

FEATURES  
source

1..3326  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="MGC:42319 IMAGE:4817835"  
/tissue\_type="Brain, hippocampus"  
/clone\_lib="NIH\_MGC\_95"  
/lab\_host="DH10B"  
/notes="vector: pbluescript"  
1..3326  
/gene="RAB15"  
/db\_xref="LocusID:376267"  
5..631  
/codon\_start=1  
/product="Ras-related protein Rab-15"  
/protein\_id="AAH40679.2"  
/db\_xref="GI:34783347"  
/db\_xref="LocusID:376267"  
/translators="MAKQVDLPRLLILIGDSGVKTCLLCRFTDNBSHSHISTIGVD  
FMKTEVDGIKVIQIWDAGERYQTIKQYRRAGQIFLYDISSERSYOHMKW  
VSDVDEVDATSLPGCGEGAPGKARRPGKANRKLCLPQPMWTKSTGHOKARR  
SLIGIRLMRSRNRGWEESKGSRRSMAWTSMKQVPAPTSTLKSRRV"

## gene

## CDS

misc\_feature 26..343  
/notes="RAB: Region: Rab subfamily of small GTPases"  
/db\_xref="CDD:cd00154"  
ORIGIN  
Query Match 93.0%; Score 3030.2; DB 9; Length 3326;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 3187; Conservative 0; Mismatches 8; Indels 132; Gaps 2;  
Qy 41 AGTCATGGCGAAGCAGTACGATGCTGCTCCGGCTCTGCTGATCGGGGACTCCGGGGT 100  
Db 1 AGTCATGGCGAAGCAGTACGATGCTGCTCCGGCTCTGCTGATCGGGGACTCCGGGGT 60  
Qy 101 GGGCAAGACCTGCTGCTGCGCTTCCCGCAACAGAGTTCACCTCTCGACATCTC 160  
Db 61 GGGCAAGACCTGCTGCTGCGCTTCCCGCAACAGAGTTCACCTCTCGACATCTC 120  
Qy 161 CACCATCGGTGTTGACCTTTAAGATGAAGACCATAGAGGTAGAGGCATCAAGTGC 220  
Db 121 CACCATCGGTGTTGACCTTTAAGATGAAGACCATAGAGGTAGAGGCATCAAGTGC 180  
Qy 221 ACAGATCTGGGACACTGTCAGGCGAGGAGATACAGACCATCACAAGCAGTACTATCG 280  
Db 181 ACAGATCTGGGACACTGTCAGGCGAGGAGATACAGACCATCACAAGCAGTACTATCG 240  
Qy 281 GCGGGCCAGGGGATATTTTGGTCTATGACATTAGCAGCAGCGCTCTTACCAGACAT 340  
Db 241 GCGGGCCAGGGGATATTTTGGTCTATGACATTAGCAGCAGCGCTCTTACCAGACAT 300  
Qy 341 CATGAAGTGGGTGAGTCACTGGATGA-----GTACGCCACAGAGCGCTCCAG 367  
Db 301 CATGAAGTGGGTGAGTCACTGGATGAAGAGCTCTGTGACGACAGAGCGCTCCAG 480  
Qy 368-----GTACGCCACAGAGCGCTCCAG 389  
Db 421 GAAGCTTTGCTTCCACAGCCCTGGATGAAGAGCTCTGTGACGACAGAGCGCTCCAG 480  
Qy 390 AAGATCTTATTCGGGAATAGGCTGTATGAGGAGCAGAAACGGCAGGTGGGAAGAGAGCAA 449  
Db 481 AAGATCTTATTCGGGAATAGGCTGTATGAGGAGCAGAAACGGCAGGTGGGAAGAGAGCAA 540  
Qy 450 GGCAGCAGCTGCGGAAGAGTATGCGATGGACTTCTATGAAAACAGTGCTGCACCAAC 509  
Db 541 GGCAGCAGCTGCGGAAGAGTATGCGATGGACTTCTATGAAAACAGTGCTGCACCAAC 600  
Qy 510 CTCACATTAAGAGTCACTTACGCGCTCTGACAGAGCTGTGCTGCGAGGCCCATAGGAAG 569  
Db 601 CTCACATTAAGAGTCACTTACGCGCTCTGACAGAGCTGTGCTGCGAGGCCCATAGGAAG 660  
Qy 570 GAGCTGGAAGGCTCCGGATCGGTGCGCAGCAATGAGTTGSCATCGGCAGAGCTGGAGGAG 629  
Db 661 GAGCTGGAAGGCTCCGGATCGGTGCGCAGCAATGAGTTGSCATCGGCAGAGCTGGAGGAG 720  
Qy 630 GAGGAGGCAAAACCCGAGGGCCAGGAACTCTTGAAGAACTGCTGCTGCTGAGTCTG 689  
Db 721 GAGGAGGCAAAACCCGAGGGCCAGGAACTCTTGAAGAACTGCTGCTGCTGAGTCTG 780  
Qy 690 TGTGGGGCAGCCCCACACGACACCCCTCTTCCCTCAGAGGCGCGTGGGAGAGAGGGGAG 749  
Db 781 TGTGGGGCAGCCCCACACGACACCCCTCTTCCCTCAGAGGCGCGTGGGAGAGAGGGGAG 840  
Qy 750 CCGGGGCTTTGGCCCTGCTGCTGCTCTCTGCTGATGACCCCTTATGAGTATCAGTAGCCA 809  
Db 841 CCGGGGCTTTGGCCCTGCTGCTGCTCTCTGCTGATGACCCCTTATGAGTATCAGTAGCCA 900  
Qy 810 CTACTCCCCCTGCTGGCCCTGAGAGCGGCTCTGCTGCTATCTCAAGACAGCCCTGCTCCC 869  
Db 901 CTACTCCCCCTGCTGGCCCTGAGAGCGGCTCTGCTGCTATCTCAAGACAGCCCTGCTCCC 960

QY 870 CAGCCGTCACCCCTGAGTGGTCTTCTTCTCAGCCCTGTTTCCCAAGCCACAGGCTGTCTAC 929  
DB 961 CAGCCGTCACCCCTGAGTGGTCTTCTTCTCAGCCCTGTTTCCCAAGCCACAGGCTGTCTAC 1020  
QY 930 GACCCCAAGATGCGCGAAGCACTGCTCTCACCATCCGACCCACCAAGCAAGCAAGCA 989  
DB 1021 GACCCCAAGATGCGCGAAGCACTGCTCTCACCATCCGACCCACCAAGCAAGCAAGCA 1080  
QY 990 GGGCTGAGTCCAGGCGCACTTTTCTCAGCTGCTCTTCTTCTCAGTGTCTCTTCTCTG 1049  
DB 1081 GGGCTGAGTCCAGGCGCACTTTTCTCAGCTGCTCTTCTTCTCAGTGTCTCTTCTCTG 1140  
QY 1050 CTTTTCTCTCTTCCCACTCTCTCTTCTCTGACCCCTCCCTCCGCTGCTTCTCTCTG 1109  
DB 1141 CTTTTCTCTCTTCCCACTCTCTCTTCTCTGACCCCTCCCTCCGCTGCTTCTCTCTG 1200  
QY 1110 CAAAGCTCTCAAAACCCCGTCCCGTGTCTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCT 1169  
DB 1201 C-AAGCTCTCAAAACCCCGTCCCGTGTCTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCT 1259  
QY 1170 CTTCTTAAGCTATCCAAAGGGATGGAACCCAGGCTCGTGGGAGGTTCCACCCTTGGATCC 1229  
DB 1260 CTTCTTAAGCTATCCAAAGGGATGGAACCCAGGCTCGTGGGAGGTTCCACCCTTGGATCC 1319  
QY 1230 AGGAAGAACCTTCCACCCTGCTCTGCTGGGTGGGCAAGGCTACAGGGTGTCTTCTCTC 1289  
DB 1320 AGGAAGAACCTTCCACCCTGCTCTGCTGGGTGGGCAAGGCTACAGGGTGTCTTCTCTC 1379  
QY 1290 TTCCCCCAACCCCTGCTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGT 1349  
DB 1380 TTCCCCCAACCCCTGCTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGT 1439  
QY 1350 GTGAGCATCGAGTGTAGGAGGAACAGCAACCGGGAGTCTCTCGAGCTTGGGCTGCC 1409  
DB 1440 GTGAGCATCGAGTGTAGGAGGAACCGCAACAGGAGTCTCTCGAGCTTGGGCTGCC 1499  
QY 1410 TACCTCTACCCATTTCCCGACCCAGAGCTTTGCTCTGCTGTGCTGTGCTGTGCTGTGCT 1469  
DB 1500 TACCTCTACCCATTTCCCGACCCAGAGCTTTGCTCTGCTGTGCTGTGCTGTGCTGTGCT 1559  
QY 1470 GGGAACTGAGCTCAGAGGCGAGTGTCTCAGAGAGGAACAAATAGGGGTGGCAGG 1529  
DB 1560 GGGAACTGAGCTCAGAGGCGAGTGTCTCAGAGAGGAACAAATAGGGGTGGCAGG 1619  
QY 1530 ATAAAGTCACTCTCTACCTCCATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1589  
DB 1620 ATAAAGTCACTCTCTACCTCCATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1679  
QY 1590 GGGTCCCAAAATTTAAAGATGTGAACCAAGGCTGTGGGTACTCCAGGGGCAAGGAGCC 1649  
DB 1680 GGGTCCCAAAATTTAAAGATGTGAACCAAGGCTGTGGGTACTCCAGGGGCAAGGAGCC 1739  
QY 1650 CTGGGTCAGTGACACTGTGAGGCAACCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1709  
DB 1740 CTGGGTCAGTGACACTGTGAGGCAACCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1799  
QY 1710 GAAGCACTAGTCTCTATGTATCAGTTAAGCAAGGAGCTGGCCAGGACAGCAGT 1769  
DB 1800 GAAGCACTAGTCTCTATGTATCAGTTAAGCAAGGAGCTGGCCAGGACAGCAGT 1859  
QY 1770 TTGCACAGCAGAGGGGAATGTAGCAACAGCAGGCGCTCTTACGGCCCATCTTCCATTTCT 1829  
DB 1860 TTGCACAGCAGAGGGGAATGTAGCAACAGCAGGCGCTCTTACGGCCCATCTTCCATTTCT 1919  
QY 1830 TAGGTAAAGAGCAATTTCTCAGATCCCAAGCGGAGGACTGAGCCTTAGCCTTTCAGCAA 1889  
DB 1920 TAGGTAAAGAGCAATTTCTCAGATCCCAAGCGGAGGACTGAGCCTTAGCCTTTCAGCAA 1979  
QY 1890 CCAAGGTTCTCTGGGACCCCAAGTTTATGGGAAGGGCAAGCACTTCATGGGAAGAGA 1949  
DB 1980 CCAAGGTTCTCTGGGACCCCAAGTTTATGGGAAGGGCAAGCACTTCATGGGAAGAGA 2039  
QY 1950 GAAGGAAGGCCCTGGGTAGAAACGCTTGGTGTCTTCTTCTTGGCCTTTTAAAGCAAGCG 2009

DB 2040 GAAGGAAGGCCCTGGGTAGAAACGCTTGGTGTCTTCTTCTTGGCCTTTTAAAGCAAGCG 2099  
QY 2010 CTATCTTTGGCCCTTCTCTCTCTGATAGGCTTGGGGTTTGGCAACCACTGTGGCTACA 2069  
DB 2100 CTATCTTTGGCCCTTCTCTCTCTGATAGGCTTGGGGTTTGGCAACCACTGTGGCTACA 2159  
QY 2070 GGTGGAGGGAAGGAGCTCTCTCTCTCAGAGTGTCTATGTTTCAAGAGTTTCTTTAAACCC 2129  
DB 2160 GGTGGAGGGAAGGAGCTCTCTCTCTCAGAGTGTCTATGTTTCAAGAGTTTCTTTAAACCC 2219  
QY 2130 ATATGGCCCAAGAGTAGCTCTGAGGAGCCCTTTTAAAGCAAGCAAGTAAATTTACCACT 2189  
DB 2220 ATATGGCCCAAGAGTAGCTCTGAGGAGCCCTTTTAAAGCAAGCAAGTAAATTTACCACT 2279  
QY 2190 TCTACTGGGGTCTCTGCTCCCAAGCTGCGTGGGCGAGGCTAGGAAGAGGGTCACTCT 2249  
DB 2280 TCTACTGGGGTCTCTGCTCCCAAGCTGCGTGGGCGAGGCTAGGAAGAGGGTCACTCT 2339  
QY 2250 TAAAGCCACATATAGCTGCACTGCTGGCTGCGAGCCAAACAAAGAACTTGGGTGTGAGT 2309  
DB 2340 TAAAGCCACATATAGCTGCACTGCTGGCTGCGAGCCAAACAAAGAACTTGGGTGTGAGT 2399  
QY 2310 ATTCACTCAACTAAGAACCAAAATCCAGGGCACTCATATGTGAAGGATAAGAACCTCACTT 2369  
DB 2400 ATTCACTCAACTAAGAACCAAAATCCAGGGCACTCATATGTGAAGGATAAGAACCTCACTT 2459  
QY 2370 CTTTACTCTCTCCAAAGAGTGGGAAAGAACCACTCAAACTTTCTCTGACTTACCA 2429  
DB 2460 CTTTACTCTCTCCAAAGAGTGGGAAAGAACCACTCAAACTTTCTCTGACTTACCA 2519  
QY 2430 AACCAAGAAACAGCAGGAGGCTGCTCAGGACTTAGGACAGGGTATAGCTTAGATG 2489  
DB 2520 AACCAAGAAACAGCAGGAGGCTGCTCAGGACTTAGGACAGGGTATAGCTTAGATG 2579  
QY 2490 GTGGAAGCAAGAGAGCAGGAAGTTGTAAATCACTGGCTTAATGAGAAAGAGAGACAGC 2549  
DB 2580 GTGGAAGCAAGAGAGCAGGAAGTTGTAAATCACTGGCTTAATGAGAAAGAGAGACAGC 2639  
QY 2550 TAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTTCAAGATGGGACTCTCTG 2609  
DB 2640 TAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTTCAAGATGGGACTCTCTG 2699  
QY 2610 GGTATCAAGACCTATGCAACATCACTGCGGCTAGGGAAGTAGGTATGTCAGCCCTCA 2669  
DB 2700 GGTATCAAGACCTATGCAACATCACTGCGGCTAGGGAAGTAGGTATGTCAGCCCTCA 2759  
QY 2670 AGTCTGTCTTCAAGCAGGACTTGAAGTTATATGGGCAAGTGGCTCCCAATCTGTGAC 2729  
DB 2760 AGTCTGTCTTCAAGCAGGACTTGAAGTTATATGGGCAAGTGGCTCCCAATCTGTGAC 2819  
QY 2730 CAGTATTTTCTAGCTTCTTCAAGATCAGGAGGCTGCAATCTGTTCTTCTCTCTAG 2789  
DB 2820 CAGTATTTTCTAGCTTCTTCAAGATCAGGAGGCTGCAATCTGTTCTTCTCTCTAG 2879  
QY 2790 CCCCCTCAGGAAGAGGACTATATTTGTACTGTACTAGGCTTCTGGAAGGGAAC 2849  
DB 2880 CCCCCTCAGGAAGAGGACTATATTTGTACTGTACTAGGCTTCTGGAAGGGAAC 2939  
QY 2850 ATGGAATCAGGATTTCTATAGCTGATAGGCTTATCCAAAGGGCCATGACTGGGAAG 2909  
DB 2940 ATGGAATCAGGATTTCTATAGCTGATAGGCTTATCCAAAGGGCCATGACTGGGAAG 2999  
QY 2910 GTATGGGAGCAGGAAGGAGATTTAGGCTGAGCTACGCTCACTCACTTAACTTTT 2969  
DB 3000 GTATGGGAGCAGGAAGGAGATTTAGGCTGAGCTACGCTCACTCACTTAACTTTT 3059  
QY 2970 GGTGGCTTGGGCACTGTCTGAGGCCAGACTGTGTTAAAGCAGGCTCTGTGCGCTGTGTAC 3029  
DB 3060 GGTGGCTTGGGCACTGTCTGAGGCCAGACTGTGTTAAAGCAGGCTCTGTGCGCTGTGTAC 3119  
QY 3030 TCGTCAACCTCTGCACTGTCTGTGAGACTTCCATCCAGGCCCGGCAAGCAGCAGCTG 3089

```

3120 TCGTCACCACTCTGCACCTGTGTCTTGAGACTCCACGAGCCACGCCACCTG 317
3090 CTCCTGAGCCTCACATATCTCCCTGTGACGGGTGAACCTTCGTCTACTGTCTCGGGTCC 3149
3180 CTCCTGAGCCTCACATATCTCCCTGTGACGGGTGAACCTTCGTCTACTGTCTCGGGTCC 3239
3150 ATATATGAATTGTGACGAGGTTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGAT 3209
3240 ATATATGAATTGTGACGAGGTTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGAT 3299
3210 TATTTCAAACCAACCAAAAAA 3236
3300 TATTTCAAACCAAAAAA 3326

RESULT 2
CNS01DX4      190517 bp      DNA      linear      PRI 17-MAY-2001
LOCUS      Human chromosome 14 DNA sequence BAC R-840119 of library RPCI-11
DEFINITION      from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION      AL139022
VERSION      AL139022.4      GI:14148954
KEYWORDS      HTG.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens (human)
REFERENCE      1 (bases 1 to 190517)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
      1 (bases 1 to 190517)
      Heilig, R., Pett, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
      Brottier, P., Cottolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
      Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,
      Gypay, G., Saurin, W. and Weissenbach, J.
      Sequencing of the human chromosome 14
      Unpublished
      2 (bases 1 to 190517)
      Genoscope.
      Direct Submission
      Submitted (17-MAY-2001) Genoscope - Centre National de Sequencing :
      BP 191 91006 Evry cedex - FRANCE (E-mail : secref@genoscope.cns.fr)
      - Web : www.genoscope.cns.fr
      On May 19, 2001 this sequence version replaced gi:10798506.

COMMENT
----- Genome Center
Center: Genoscope / Centre National de Sequencing
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : C-2503G16 (AC-AL355076)
Downstream BAC (overlapping the SP6 end) : C-2300N4 -----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.14x in Q20 bases; sum-of-contigs -----

Overall quality chart :
Range : bases
0 :
1 - 9 :
10 - 19 :
20 - 29 : 45
30 - 39 : 444
40 - 49 : 4501
50 - 59 : 11521
60 - 69 : 12036
70 - 79 : 23986
80 - 89 : 56256
90 - 99 : 81728
-----
Percentage of bases with a quality value >= 40 : 99 %

```

FEATURES	source	Location/Qualifiers	ORIGIN	Query Match	Score	DB	Length
Percentage of bases with a quality value >= 40 : 99 %	1. .190517			82.8%	2696.6	9	190517
		/organism="Homo sapiens"					



Best Local Similarity 98.8%; Pred. No. 0; Matches 2717; Conservative 0; Mismatches 34; Indels 0; Gaps 0;			
Qy	473	TGGCATGGACTTCTATGAACAAGTGCCTGCGACCAACCTCAACATTTAAAGAGTCATTAC	532
Db	157619	TGCCAGGGAGAGGAGACATGCACTAACTGTGGCCCTTTGGTTCCAGTCATTAC	157678
Qy	533	GGCTCTGACAGACTGGTGTGTGAGGCCCATAGAAAGAGAGCTGGAAAGGCTCCCGATGG	592
Db	157679	GGCTCTGACAGAGCTGGTGTGTGAGGCCCATAGAAAGAGAGCTGGAAAGGCTCCCGATGG	157738
Qy	593	TGCCAGCAATGAGTTGGCACTGTCAGAGCTGGAGGAGGAGGGCAACCCGAGGGGCC	652
Db	157739	TGCCAGCAATGAGTTGGCACTGTCAGAGCTGGAGGAGGAGGGCAACCCGAGGGGCC	157798
Qy	653	AGCGAACTCTTCGAAAACCTGCTGGTGTGCTGAGTCCTGTGGGGCAACCCCAACGACAC	712
Db	157799	AGCGAACTCTTCGAAAACCTGCTGGTGTGCTGAGTCCTGTGGGGCAACCCCAACGACAC	157858
Qy	713	CTCTTCCCTCAGAGAGCCGTCGGGAGAGGGAGAGGGAGAGGGGCTTGGCCCTGCTGTGT	772
Db	157859	CTCTTCCCTCAGAGAGCCGTCGGGAGAGGGAGAGGGGCTTGGCCCTGCTGTGTGT	157918
Qy	773	CCCTGCTGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCCCCCTGCTGGCCCTGA	832
Db	157919	CCCTGCTGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCCCCCTGCTGGCCCTGA	157978
Qy	833	GAGCGGCTGCTGTCTATCTCAAGCAGCCCTGTGCCAGCCCGTCCACCTGGAGTGT	892
Db	157979	GAGCGGCTGCTGTCTATCTCAAGCAGCCCTGTGCCAGCCCGTCCACCTGGAGTGT	158038
Qy	893	CTTCTTTCAGGCTGTTCCTCCAGCCACAGGCGCTGCTACGACCCCAAGATGTCGCAAGC	952
Db	158039	CTTCTTTCAGGCTGTTCCTCCAGCCACAGGCGCTGCTACGACCCCAAGATGTCGCAAGC	158098
Qy	953	ACTGTCTCACATCCCGCACCCACAGACAAAGCCAGGCGCTGCTACGACCCCAAGATGTC	1012
Db	158099	ACTGTCTCACATCCCGCACCCACAGACAAAGCCAGGCGCTGCTACGACCCCAAGATGTC	158158
Qy	1013	AGCTGCTCTTCTCGTGATGATGCTGCTCTCTCTGCTTTTCTCTCTCTCTCTCTCTCTCT	1072
Db	158159	AGCTGCTCTTCTCGTGATGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	158218
Qy	1073	TCCTTCTCTGACCCCTCCCTCCGCTGCTTTCGATCAAGCTCTCTCAACCCCGTCCC	1132
Db	158219	TCCTTCTCTGACCCCTCCCTCCGCTGCTTTCGATCAAGCTCTCTCAACCCCGTCCC	158278
Qy	1133	CGGTGTCTCTGTGTGAGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1192
Db	158279	CGGTGTCTCTGTGTGAGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	158338
Qy	1193	GGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAGAACCTCTCAACCTGCTCT	1252
Db	158339	GGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAGAACCTCTCAACCTGCTCT	158398
Qy	1253	CGTGGGTGGGCCAAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1312
Db	158399	CGTGGGTGGGCCAAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	158458
Qy	1313	ATGTGCAATGGGCTGCTCCCGAGTGAAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1372
Db	158459	ATGTGCAATGGGCTGCTCCCGAGTGAAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	158518
Qy	1373	AACAGCAACCGGGAGTCTCGAGCTGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1432
Db	158519	AACAGCAACCGGGAGTCTCGAGCTGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	158578
Qy	1433	GAGCTTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1492
Db	158579	GAGCTTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	158638
Qy	1493	TGCTTTCAGAGAGGAAACAAATGAGGGGTGGCAGGGATGAGGAGTAAAGTCACTCTCTCT	1552













Db 7622 TTCTCTCAGCCTAACTCCAAATTTAAATAATATAACCAAAACCTATAATATCTCCAAA 7563  
Qy 1638 GCAAGGAGAGCCTGGGTGTCAGTGAACCTGTGAGGCAACCATGCACTCCACAAAGGGA 1697  
Db 7562 AAAAAAACCCTAAAATCAATAACACTATCAACCAACCATACACTCCACAAAABAAA 7503  
Qy 1698 GCATTTGGAATGAAGACTAGCTCTCTATGTATCATCAGGTATAGACGAAGGAGCTGGCC 1757  
Db 7502 ACATTTAAAAATAAAAAAATAAATTTAAATTTAAAAAATAAAAAAATAAATAAACC 7443  
Qy 1758 AGGACAGCAGTGTTCAGCAGCAGAGGGAATGTAGCAACGAGGCGCTCTCTAGGCCCA 1817  
Db 7442 AAAAACAACATTTTACACAAACAAAAAATAATATAACCAACCAAAACCTCTTAAACCCCA 7383  
Qy 1818 TCTTCCATTTCTTAGTGAAGAGCAATTTCTCAGACTCCAGGCGGAGGACTGAGCCT 1877  
Db 7382 TCTTCCATTTCTTAAATATAAAAAAACAATTTCTCAAACTCCCAACCGAATAAACA 7323  
Qy 1878 AGCCTTCAGCAACCAAGGTTCTCTGGGACCCAAAGTTTATGGGAGAGGCAAGACTT 1937  
Db 7322 AACCTTCAACCAACCAAAATTTCTCTAAACCCCAAAATTTATAAAAAAATAAAAAA 7263  
Qy 1938 CATGGGAGAGAGAGAGGAGGCTGGGTAGAAAGCTGTGGTCTCTCTTTGGCCTT 1997  
Db 7262 CATAAAAAATAAAAAAATAAAAAAACCCTAAATAAAAGCTTAACTATCTCTTTAACCTT 7203  
Qy 1998 TAAGACAAAGCGCTCATCTTCCCTCTACCTCTGATAGGCTTCAGGCTTGGCCAAAC 2057  
Db 7202 TAAACAAAGCTCATCTTACCTCTACCTCTCTTAACTTAACTTAACTTAACTTAA 7143  
Qy 2058 ACTGTGCTACAGGTGAGGAGAGGAGCTCTCTCTCCAGAGTGTCTATGTTTCAAGAGT 2117  
Db 7142 ACTATAACTACAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 7083  
Qy 2118 TTCTTTTAAACCCATATGCCCCAGAGTAGCTCTGAGGAGGCTTTTAAAGCGAAACAG 2177  
Db 7082 TTCTTTTAAACCCATATATACCCAAATAATACTCGTAAAAAACCCTTTTAAACCGAA 7023  
Qy 2178 TAAATTTACCACTTCTACTGGGGTTCCTGCCCAACCGTCCCAAGGTGGCGAGGCTTAGAA 2237  
Db 7022 TAAATTTACCAATTTCTTAAATTTCTTACCCACCGTCCCAATAATAAGGAACTTAA 6963  
Qy 2238 GAGGTCATTTTAAAGCCACATTAAGTGTGCTGCTGGGTGAGGCTTAAAGCAAGAAC 2297  
Db 6962 AAAAAATCATTTCTTAAACCAACATTAATACTACCTAGGTAACTCAACCAACCAAA 6903  
Qy 2298 TGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGCACTCATATGTGAAGATA 2357  
Db 6902 TAAATATTAATAATTTCTCACTAACTAAAAACCAAAATCCAAACACATCATATA 6843  
Qy 2358 AGAACCTCACCTTCTTACTCTCTCAAAAGAGTGGGGAAGAACCACTCAAACTTTCT 2417  
Db 6842 AAACCTCACCTTCTTACTCTCTCAAAAGAGTGGGGAAGAACCACTCAAACTTTCT 6783  
Qy 2418 CCTGACTTACCAACCAAGGAAACAGCAGGAGGAGGCTGCTGAGGCTTAGGCAAGG 2477  
Db 6782 CCTAATTTACCAACCAAGGAAACCAACCAAAATAAATACTCAAACTTAAACCAAA 6723  
Qy 2478 ATAGCTTAGTGGTGAAGCAAGAGGAGGAGGAGGAGTGTGAATCACTGGCTTAATGA 2537  
Db 6722 ATAACTTTAAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 6663  
Qy 2538 AAAGGAGAGCTAACTTCTAGATGAAGTGTGAGTGGGAGGCTTCTTCTTGAAGA 2597  
Db 6662 AAAAAAACAACCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6603  
Qy 2598 TGGGACTCTTTGGGTATCAAGACCTATGCCACATCACTGGGCTTAGGGAAGTAGTGA 2657  
Db 6602 TAAACCTCTTAAATATCAAAACCTATACCAATCACTCAAACTTAAATAAATAA 6543  
Qy 2658 TGGCAGCCCTCAAGCTGTCTTACGACGAGGAGCTTGAAGAGTATATTTGGGCACTG 2717  
Db 6542 TACCAACCTCAATCTATCTTCAACCAAAACCTTAAATAAATAAATAAATAAATAA 6483

Qy 2718 CAATCTGTGGACCAAGTATTTTCACTTTTCCCTGAAGATCAGCAGGGTGCATTCATTGTC 2777  
Db 6482 CAATCTATAAACCAATATTTTCACTTTTCCCTGAAGATCAGCAGGGTGCATTCATTATC 6423  
Qy 2778 TTTTCTCTCTAGCCCTCAGGAAAGAGCACTATTTTGTACTGTACCTTACGGGTTCT 2837  
Db 6422 TTTTCTCTCTAGCCCTCAGGAAAGAGCACTATTTTGTACTGTACCTTACGGGTTCT 6363  
Qy 2838 GGAAGGAAACATGGAATCAGGATTTCTATAGACTGTAGGCTTATCCCAAGGGCAT 2897  
Db 6362 AAAAAAACAACATTAATAATCTTAATACTTAATACTTAATACTTAATACTTAATA 6303  
Qy 2898 GACTGGGAAAGGTATGGGAGAGGAGATTTGGGATTTTGGGTCAGCTACGCTCA 2957  
Db 6302 AACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6243  
Qy 2958 CCCTAAACTTTTGGTGGCCCTGGGCGATGCTTGTAGGCGGCGAGCTGTTAAGCAGGCTG 3017  
Db 6242 CCCTAAACTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6183  
Qy 3018 TGGCCTGTCTTACTGCTCAGCAGCTCTGCACTGCTGCTTGTAGACTCCATCCAGCC 3077  
Db 6182 TAACCTATTTTACTGCTCAGCAGCTCTTAACTTAACTTAACTTAACTTAACTTAA 6123  
Qy 3078 GCAGCCACCTGCTCTGAGCCTCCACTATCTCTCTGCTGAGGCTGAACCTTCTGTG 3137  
Db 6122 ACAGCCACCTGCTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 6063  
Qy 3138 TGCTCGGTCCTATATGAAATTTGTAGGAGGCTTCTATTTTAAACACAGATGTTTA 3197  
Db 6062 TATCTGAACTTCAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6003  
Qy 3198 CAAATAAAGATTTTCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3257  
Db 6002 CAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5943

RESULT 8  
AX399903  
LOCUS AX399903 1054 bp DNA linear PAT 06-JUN-2002  
DEFINITION Sequence 74 from Patent WO0218424.  
ACCESSION AX399903  
VERSION AX399903.1 GI:21336157  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Tang, Y.T., Asundi, V., Zhou, P., Xue, A.J., Ren, F., Zhang, J.,  
Wang, J.R., Zhao, Q.A., Wang, D., Liu, C., Drmanac, R.T. and Wehrman, T.  
Nucleic acids and polypeptides  
TITLE Patent: WO 0218424-A 74 07-MAR-2002;  
JOURNAL HYSEQ, INC. (US)  
FEATURES  
Location/Qualifiers  
source 1..1054  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
81..719  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD33347.1"  
/db\_xref="GI:21336158"  
/db\_xref="RENTREMBL:CAD33347"  
/translation="MAKQYDVLFRLLIGDSGVKTCCLCRFTDNFPHSHSTIGVD  
FKMTIEVDGIVKRIQIWDVTAQERYOTIKQYVRAQGIPLVYDIESSEYOHMKW  
VSDVDEYAPGVQKILIGNKADKEQVGEQOQAKKEYGMDFFYTSACTNLNIKE  
SFTRLTELVLQAHKRLLEGURMARNALALALEEEKPEGPANSSKTCWC"

CDS  
Query Match 30.7%; Score 1000.2; DB 6; Length 1054;  
ORIGIN



Best Local Similarity 99.6%; Pred. No. 2.3e-239; Matches 1002; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
Qy	2	GCCCGCTGCCGCGCCGAGTTC	CGCGCCGCTGCGCCCGAGTACGA	61					
Db	38	GCCCGCTGCCGCGCCGAGTTC	CGCGCCGCTGCGCCCGAGTACGA	97					
Qy	62	TGTCCTGTCGCGCTGCTGAT	CGCGGCTGCGCGAGACCTGCGTGTG	121					
Db	98	TGTCCTGTCGCGCTGCTGAT	CGCGGCTGCGCGAGACCTGCGTGTG	157					
Qy	122	CCGCTTCAACGACAAAGTTC	CACTCTCGCACATCTCCACCATCGTGTTGACTTTAA	181					
Db	158	CCGCTTCAACGACAAAGTTC	CACTCTCGCACATCTCCACCATCGTGTTGACTTTAA	217					
Qy	182	GATGAAGACCATAGAGTAG	ACGCGCATCAAGTGGGATACAGATCTGGGACACTGACG	241					
Db	218	GATGAAGACCATAGAGTAG	ACGCGCATCAAGTGGGATACAGATCTGGGACACTGACG	277					
Qy	242	GCAGGAGATACAGACCAT	CAAAAGCAGTACTATCGCGGCGCCAGGGGATATTTT	301					
Db	278	GCAGGAGATACAGACCAT	CAAAAGCAGTACTATCGCGGCGCCAGGGGATATTTT	337					
Qy	302	GGTCTATGACATTAGCAG	CGCGCTCTTACAGCACATCATGAAGTGGGTGAGTACGT	361					
Db	338	GGTCTATGACATTAGCAG	CGCGCTCTTACAGCACATCATGAAGTGGGTGAGTACGT	397					
Qy	362	GGATGATGCGCACCAAG	CGCTCCAGAGATCTTATTTGGGATGAAGTATGAGGA	421					
Db	398	GGATGATGCGCACCAAG	CGCTCCAGAGATCTTATTTGGGATGAAGTATGAGGA	457					
Qy	422	GCAGAAACGCGCAGTGG	GAAGAGAGCAGCAGCTGGCGAGGAGTATGGCATGGA	481					
Db	458	GCAGAAACGCGCAGTGG	GAAGAGAGCAGCAGCTGGCGAGGAGTATGGCATGGA	517					
Qy	482	CTTCTATGAACAAAGT	CGCTGCACCAACCTCAACATTAAAGAGTCATTACGCGCTTGAC	541					
Db	518	CTTCTATGAACAAAGT	CGCTGCACCAACCTCAACATTAAAGAGTCATTACGCGCTTGAC	577					
Qy	542	AGAGCTGGTCTGCAG	CGCCCATAGGAGAGCTGGAGGCTCCGGATGCTGCGACAA	601					
Db	578	AGAGCTGGTCTGCAG	CGCCCATAGGAGAGCTGGAGGCTCCGGATGCTGCGACAA	637					
Qy	602	TGAGTTGGCACTGGC	AGAGCTGGAGGAGGAGGCAAAACCGAGGGCCAGCGAACTC	661					
Db	638	TGAGTTGGCACTGGC	AGAGCTGGAGGAGGAGGCAAAACCGAGGGCCAGCGAACTC	697					
Qy	662	TTGGAACCTGCTGCT	GAGTCTGTGTGTGGGCAACCCACACGACACCCCTCTTCCC	721					
Db	698	TTGGAACCTGCTGCT	GAGTCTGTGTGTGGGCAACCCACACGACACCCCTCTTCCC	757					
Qy	722	TCAGGAGCCGCTGG	CAGACAGGGAGCGGGCTTTGCGCTGCTCTCTGCTG	781					
Db	758	TCAGGAGCCGCTGG	CAGACAGGGAGCGGGCTTTGCGCTGCTCTCTCTGCTG	817					
Qy	782	TGATGACCTTATG	ATGATCAGTAGCCACTACTCCCGCTGCTGGCCCTCAGAGCGGCTC	841					
Db	818	TGATGACCTTATG	ATGATCAGTAGCCACTACTCCCGCTGCTGGCCCTCAGAGCGGCTC	877					
Qy	842	TGCTGTCACTCA	AGACGCGCTGTCTCCAGCGCGTCCACCTGGAGTGTCTTCTTCAG	901					
Db	878	TGCTGTCACTCA	AGACGCGCTGTCTCCAGCGCGTCCACCTGGAGTGTCTTCTTCAG	937					
Qy	902	CCGTGTTTCCCG	CGCCAGCGCTGTACGACCCCGCATGTCGCGCAAGCACTGCTCA	961					
Db	938	CCGTGTTTCCCG	CGCCAGCGCTGTACGACCCCGCATGTCGCGCAAGCACTGCTCA	997					
Qy	962	CCATCCCGCAC	CCACAGCAACAGCGCGCTGGAGTCCAGGCCA 1007						
Db	998	CCATCCCGCAC	CCACAGCAACAGCGCGCTGGAGTCCAGGCCA 1043						

```

Qy 3018 TGGCTGTTTACTCGTCACCACTCTGACACCTGCTGCTTTGAGACTCCATCCAGCCCGAG 3077
Db 673 TGGCTGTTTACTCGTCACCACTCTGACACCTGCTGCTTTGAGACTCCATCCAGCCCGAG 732
Qy 3078 GCAGCCACCTGCTCCTGAGCTCCACTATCTCCTGTGACGGGTGAACCTTCGTGACTG 3137
Db 733 GCAGCCACCTGCTCCTGAGCTCCACTATCTCCTGTGACGGGTGAACCTTCGTGACTG 792
Qy 3138 TGTCTCGGGTCCATATATGATTGTGACAGGGTTCATCTATTATTTAAACACAGATGTTTA 3197
Db 793 TGTCTCGGGTCCATATATGATTGTGACAGGGTTCATCTATTATTTAAACACAGATGTTTA 852
Qy 3198 CAAATAAAGATTATTTCAACACCAACCAAAAAAAAAAAAAAAAA 3240
Db 853 CAAATAAAGATTATTTCAACACCAACCAAAAAAAAAAAAAAAAA 895

RESULT 10
BD205055 895 bp DNA linear PAT 17-JUL-2003
LOCUS Human nucleic acid sequence originating in cystic cancer tissue.
DEFINITION BD205055
ACCESSION BD205055.1 GI:33014825
VERSION JP 2002512023-A/9.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Specht, T., Hinzmann, B., Schmitt, A., Pilarsky, C., Dahl, E. and
Rosenthal, A.
Human nucleic acid sequence originating in cystic cancer tissue
Patent: JP 2002512023-A 9 23-APR-2002;
METAGEN GESELTSCHAFT FUER GENOM FORSCHUNG MBH
OS Homo sapiens (human)
PN JP 2002512023-A/9
PD 23-APR-2002
PR 15-APR-1999 JP 2000544779
PR 21-APR-1998 DE 198 18 619, 3
PI THOMAS SPECHT, BERND HINZMANN, ARMIN SCHMITT, CHRISTIAN PILARSKY,
PI EDGAR DAHL,
PI ANDRE ROSENTHAL
PC C12N15/09, A61K38/00, A61K39/395, A61K39/395, A61K48/00, A61P13/10,
PC A61P35/00,
PC C07K14/47, C07K16/18, C12N5/10, C12P21/02, C12P21/08, C12Q1/68, PC
C12N15/00,
PC A61K37/02, C12N5/00
CC Human nucleic acid sequence originating in cystic cancer CC
tissue
FH Key Location/Qualifiers
FT source 1..895
FT Location/Qualifiers
1..895
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="caxon:9606"

FEATURES
source
Query Match 26.6%; Score 866.2; DB 6; Length 895;
Best Local Similarity 99.5%; Pred. No. 8.9e-206;
Matches 879; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 2358 AGAACCTCAGTCTTCTTCTCTCCAAAAGAGTGGGAAAGAACCATCAACCTTTCTT 2417
Db 14 AGAACCTCAGTCTTCTTCTCTCCAAAAGAGTGGGAAAGAACCATCAACCTTTCTT 73
Qy 2418 CCTGACTTACCAACAGGAAACAGGAGGGTGGCTCAGGACTAGGACAGGGT 2477
Db 74 CCTGACTTACCAACAGGAAACAGGAGGGTGGCTCAGGACTAGGACAGGGT 133
Qy 2478 ATAGCTTAGTGGTGAAGCAAGGAGACGAGGAGGAGTGTAAATCACTGGCTATGAGA 2537

```

```

Db 134 ATAGCTTAGATGTGGAAAGCAAGGAGACAGGAAAGTTGTAATCACTGGCTAATGAGA 193
Qy 2538 AAAGGACACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTGTGCTTCTTTGAAGA 2597
Db 194 AAAGGACACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTGTGCTTCTTTGAAGA 253
Qy 2598 TGGGACTCCTTTGGGTATCAAGACCTATGCCACATCACTGCGGCTAGGGAAGTAGGTGA 2657
Db 254 TGGGACTCCTTTGGGTATCAAGACCTATGCCACATCACTGCGGCTAGGGAAGTAGGTGA 313
Qy 2658 TGGCAGCCCTCAAGTCTGCTTCCAGCCAGGACTTTGAGAAAGTTATATTGGGACGTGGCTC 2717
Db 314 TGGCAGCCCTCAAGTCTGCTTCCAGCCAGGACTTTGAGAAAGTTATATTGGGACGTGGCTC 373
Qy 2718 CAATCTCTGGACCAAGTATTTTTCAGCTTTCCCTCAAGATCAGGCGGTGCCATTCATTGTC 2777
Db 374 CAATCTCTGGACCAAGTATTTTTCAGCTTTCCCTCAAGATCAGGCGGTGCCATTCATTGTC 433
Qy 2778 TTTCTCTCTAGCCCTCCTCAGGAAAGGACTATATTGTTACTGTGACCTTAGGGTTCT 2837
Db 434 TTTCTCTCTAGCCCTCCTCAGGAAAGGACTATATTGTTACTGTGACCTTAGGGTTCT 493
Qy 2838 GGAAGGAAACATGGAATCAGGATCTATAGACTGATAGCCCTATCCCAAGGGCCAT 2897
Db 494 GGAAGGAAACATGGAATCAGGATCTATAGACTGATAGCCCTATCCCAAGGGCCAT 553
Qy 2898 GACTGGGAAAGGATGAGGAGCAAGAGAAATTTGGGATTTTAGGCTGCAGCTACGCTCA 2957
Db 554 GACTGGGAAAGGATGAGGAGCAAGAGAAATTTGGGATTTTAGGCTGCAG-TACGCTCA 612
Qy 2958 CCCTAAACTTTTGGTGGCTGGGCACTGTCTTGAGGCCAGACTGTTAAAGCAGGCTCTGC 3017
Db 613 CCCTAAACTTTTGGTGGCTGGGCACTGTCTTGAGGCCAGACTGTTAAAGCAGGCTCTGC 672
Qy 3018 TGGCTGTTTACTGCTCACCACTCTGCACTGCTGCTTGTGAGACTCCATCCAGCCCGAG 3077
Db 673 TGGCTGTTTACTGCTCACCACTCTGCACTGCTGCTTGTGAGACTCCATCCAGCCCGAG 732
Qy 3078 GCAGCCACCTGCTCCTGAGCTCCACTATCTCCTGTGACGGGTGAACCTTCGTGACTG 3137
Db 733 GCAGCCACCTGCTCCTGAGCTCCACTATCTCCTGTGACGGGTGAACCTTCGTGACTG 792
Qy 3138 TGTCTCGGGTCCATATATGAAATGTGACAGGGTTCATCTATTATTTAAACACAGATGTTTA 3197
Db 793 TGTCTCGGGTCCATATATGAAATGTGACAGGGTTCATCTATTATTTAAACACAGATGTTTA 852
Qy 3198 CAAATAAAGATTATTTCAACACCAACCAAAAAAAAAAAAAAAAA 3240
Db 853 CAAATAAAGATTATTTCAACACCAACCAAAAAAAAAAAAAAAAA 895

RESULT 11
LOCUS BC027769 3139 bp mRNA linear ROD 11-DEC-2003
DEFINITION Mus musculus RAB15, member RAS oncogene family, mRNA (cDNA clone
MGC:38375 IMAGE:5345297), complete cds.
ACCESSION BC027769
VERSION BC027769.1 GI:20380721
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3139)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Narusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

```

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT	
				McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.					
				Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences					
				Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)					
				12477932					
				2 (bases 1 to 3139)					
				Strausberg, R.					
				Direct Submission					
				Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
				NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>					
				Contact: MGC help desk					
				Email: <a href="mailto:cgabs-i@mail.nih.gov">cgabs-i@mail.nih.gov</a>					
				Tissue Procurement: Jeffrey Green M.D.					
				cDNA Library Preparation: Life Technologies, Inc.					
				cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
				DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland,					
				Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a>					
				Contact: <a href="mailto:nisc_mgc@hghri.nih.gov">nisc_mgc@hghri.nih.gov</a>					
				Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hachighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W., Teurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.					
				Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>					
				Series: IRAC Plate: 55 Row: d Column: 5					
				This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19527265					
FEATURES				Location/Qualifiers					
source				1. 3139					
				/organism="Mus musculus"					
				/mol_type="mRNA"					
				/strain="FVB/N"					
				/db_xref="taxon:10090"					
				/clone="MGC:38375 IMAGE:5345297"					
				/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."					
				/clone_lib="NCI CGAP Mam6"					
				/lab_host="DH10B"					
				/note="Vector: pCMV-SPORT6"					
gene				1. 3139					
				/gene="Rab15"					
				/db_xref="LocusID:104886"					
				/db_xref="MGI:1916865"					
				208. 846					
CDS				/codon_start=1					
				/product="Rab15, member RAS oncogene family"					
				/protein_id="AAH27769.1"					
				/db_xref="GI:20380722"					
				/db_xref="LocusID:104886"					
				/translation="MAKQYDVLFRLLIGDSGVKTCILCRFDNEPHSHSTIGVDYFKMTIDVDGIVKVIQIWDTAGOERYOTITKQYVRAQGIPLVYDITSSRSYOHIMKVFSDYDEYAPGVQKILIGNKAEQKQVQGLAKSYGMDFTYTSACTNLNIKSFTRVLVEYQARKEUDGURTRASNLAUELEEDGKGFANSTKTCWC"					
				229. 720					
misc_feature				/note="RAB; Region: Rab subfamily of small GTPases"					

Db 1124 ACGAGGCTCAAGGCTGGAGTCAAGGCCACATTCAG-----ATTCTTTAGTGAATCTGCT 1177  
Qy 1040 CTCTTCTCTGCTTTTCTCTCTCTCCCACTTCTCTCTCTCTGACCCCTCCCTCCCGGTG 1099  
Db 1178 GGGCTCTCTGCTTTTCT 1233  
Qy 1100 GGTTCGTATCAAGCTCTCAAAACCCCGTCCCGGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1159  
Db 1234 TAAATCGCATCCCTATCATCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1293  
Qy 1160 TCTTTCT 1219  
Db 1294 TGCTACTGCACT 1351  
Qy 1220 CTTTGGATCCAGGAAGAACCTCTCAACCTGCTCTGTTGGTGGGCAAAAGGCTACAGGGTG 1279  
Db 1352 GCTCTGCCCCAGGAGGACTGTTTCAACCTG-CTCGGGGTGGGTGAGGCTA-AGGGTG 1409  
Qy 1280 CTTCT 1339  
Db 1410 CTTCT 1446  
Qy 1340 ACCTCGGAAGTGGAGCATCAGGTAGGAGGAACACAGCAACCGGGAGTCTCTGAGCCT 1399  
Db 1447 GCCTGGGAAGGTAGAGCACTCAGGTAGGAGGAACACTA--CTAGGTGGCCCTCAAGCC 1504  
Qy 1400 GGGCTGCCCCCTACCTCTACCTACCTTCCCGACAGAGCTTTGCCCTTGTCTGGCTGCCCGC 1459  
Db 1505 TGGACTGGCCCTCTCTGCTTGGCTTCTGCTGAGTCTTGGCTTGGCTTGGCTTGGCTTGGCTT 1564  
Qy 1460 CTGCTCTCTTTGGGAACCTGAGCTCAGAGGCAAGTCTCTCAGAGGAAGGAAACAAATGA-- 1517  
Db 1565 GTGCTCTGTGAGGAACCTGAGTCAGAGGCAAGGACAGCAGAGGAAGGAACGAAGAA 1624  
Qy 1518 ----GGGGTGCAGGATAAAGTACCTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1573  
Db 1625 CGATGGGGTGCAGGGAT--TAAGTCACTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1675  
Qy 1574 CAATTTCTCTCACTGCTGCCCAATTTAAAGATGTGGCAACAGGCTGTGGGTACTTCC 1633  
Db 1676 ----TCTCTCACTGATTCACAAATTTAAAGATGGGATCGAGTCTGAGGACAC-CA 1730  
Qy 1634 AGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCAAGCCCAACCATGCACTCCACAAAG 1693  
Db 1731 AAGGGCAAGGGAGCCCTGGGGTCA--GATACTGTATAGTCACTGTGTAAGTCTCAGATAG 1788  
Qy 1694 GGGAGCATTTGGAAATGAAGACTAGCTCTCTATGTATCAGGTTTAAGCAAGGAGAGCT 1753  
Db 1789 AGGCAAGAGCGGATCGGAAGGGCGGGCTCTTAGAACCCAGGGTTAGAGCAAGAGAGGG 1848  
Qy 1754 GGCCAGGGACAGAGTTTTCACAGCAGAGGGGAATGTAGCAACAGCAGGGCTCTCTAGGC 1813  
Db 1849 AGC-----TTGCGAGAACAGGAGTGTCTGTGGCACTACAGGGCATCAGAGGT 1896  
Qy 1814 CCCATCTTCTCACTTTCTTAGTAAGAAGACATTTCTCTCAGACTCCCGAGGGAGGACTGA 1873  
Db 1897 GACAAGTAATC-----AACCATCTCTCTTCACTGA 1927  
Qy 1874 GCCTAGCCTTCAGCAACCAAGGTTCTCTGGGACCCAAAGTTTATGGGAAGGGCAAG 1933  
Db 1928 ACCCAGCCTTAGGGAGTTAGGGTTTCCCGAGAGTCTGGCAACAGGGCTTAAGACAAGGG 1987  
Qy 1934 ACTTCATGGGAAGAGAGGAAGGCGCTGGGTAGAAACGCTTGGTGTCTTCTCTTTGG 1993  
Db 1988 AAGAGAAGGGTGGTA-----TGATGGAATATCTTGGTGTCTTCTGCTTCTCTCTCTCTCT 2034  
Qy 1994 CTTTAAAGCAAAAGCGCTCATCTTGGCCCTCTACCTCTCTGATAGGCTTGAGGGTTTGGCAA 2053  
Db 2035 CAG-----GACTCCCTCCCTCTGGTTCCTTATTTCTAAGTCGCGCTTGGCAA 2082  
Qy 2054 CCACACTGTGTGCTACAGGTGAGGGAAGAGGACTCTTCTCTCCAGAGTGTCTATGTTCAAG 2113  
Db 2083 CTGCTGTGTGGCTACCAAGGGAGGGAAGAGACACCCCTCTCTCCAGAGTTCTATGCTCAGG 2142

Qy 2114 AAGTTTCTTTAACCCTATATGGCCCAAGAGTAGCTCTAGGAGGCCCTTTTAAAGACGGAA 2173  
Db 2143 AAGTTTCTTTAACCCT--TGTGGTTCAAGAGTAGCTTGAAGGAGGCCCTGAAAAACAAC 2200  
Qy 2174 CAAGTAATTTA-----CCAGTTCTACTGGGGTTCCTGCCCAACCTCCCAAGTGGGGAG 2228  
Db 2201 AAACAAACAAACAAAGCCCAACCCCTGGTAGGGTTCATCTTACATTTCCCAAGTGGGAGA 2260  
Qy 2229 GCCTAGCAAGAGGTCTATTCTTAAGCCACATTTAGCTGCACTGCTGGCTGCAGCCAAA 2288  
Db 2261 GCTCAGGTAAAGAGAGTCCCAAGATGCACTTCAAGCCTGGGGCTGAGCCAAAGCAGA 2320  
Qy 2289 ACAAGAACTGGGTGT---TCAGTATTTCATCAACTAAGAACCC---AAAAATCCAGGGCACT 2342  
Db 2321 GGAAGAACTAGGCATGACCGAGCATTCAGGACCAAGGTCTTTTCAATTCAGGGCACA 2380  
Qy 2343 CATATGTGAAGGATGAAGAACT 2400  
Db 2381 GCTGTGATGACAGAGAGCCCAACCTCTTTAGTCTCTTAAAAAGGAGATGGGGGTGAA 2440  
Qy 2401 ACCATCAAACTTTCTCTCTGACTTACCAACCAAGGAAACAGCAGAGAGGTGGCTCA 2460  
Db 2441 ATCAATGGGATTTTCTCCTGACTTAGCAACAAGGAAGCAGGAGAGGAGAACAGCTCA 2500  
Qy 2461 GGACTTA 2467  
Db 2501 GGATTTA 2507

RESULT 12  
AX781147  
LOCUS AX781147 676 bp DNA linear PAT 14-JUL-2003  
DEFINITION Sequence 3304 from Patent WO03039443.  
ACCESSION AX781147  
VERSION AX781147.1 GI:32698216  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,  
Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 3304 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE);  
Ludwig-Maximilian-Universitaet Muenchen (DE);  
PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)

FEATURES  
source 1..676  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 19.4%; Score 633; DB 6; Length 676;  
Best Local Similarity 93.6%; Pred No. 2.6e-147;  
Matches 633; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 2572 CTAGGCTGGAGTTGCTCTTCTTTGAAGATGGGACTCCTTTGGGTATCAAGACCTATGCCACAT 2631  
Db 1 CTAGGCTGGAGTTGCTCTTCTTTGAAGATGGGACTCCTTTGGGTATCAAGACCTATGCCACAT 60  
Qy 2632 CACACTGGGGCTAGGAAGTAGGTATGCCGCCCTCAAGTCTGTCTTTCAGCCAGGGACT 2691  
Db 61 CACACTGGGGCTAGGAAGTAGGTATGCCGCCCTCAAGTCTGTCTTTCAGCCAGGGACT 120  
Qy 2692 TGAGAAGTTATTTGGCAGTGGCTCCAACTCTGTGGACCACTATTCAGCTTTCCTGAA 2751  
Db 121 TGAGAAGTTATTTGGCAGTGGCTCCAACTCTGTGGACCACTATTCAGCTTTCCTGAA 180  
Qy 2752 GATCAGGAGGGGTGCCATTATTGTCTTCTCTCTAGCCCCCTCAGGAAAGAGGACTA 2811

181 GATCAGCGGGTGCATTCTATGCTCTCTCTAGCCCTCCTAGGAAGAAGACTA 240  
2812 TATTGTACTGTACCTTAGGGTTCGGAAGGAAACATGGATCAGGATCTTAGAC 2871  
241 TATTGTACTGTACCTTAGGGTTCGGAAGGAAACATGGATCAGGATCTTAGAC 300  
2872 TGATAGCCCTATCCCAAGGGCCATGACCTGGGAAAGGTATGGAGCAGAAGAGATT 2931  
301 TGATAGCCCTATCCCAAGGGCCATGACCTGGGAAAGGTATGGAGCAGAAGAGATT 360  
2932 GGGATTTTAGGGTGCAGTACGCTCACCTAAACCTTTTGTGGCTGGGCACTGTCTTGA 2991  
361 GGGATTTTAGGGTGCAGTACGCTCACCTAAACCTTTTGTGGCTGGGCACTGTCTTGA 420  
2992 GGCCAGACGTGTTAAGCAGGCTGTGCTGGCTGTGTTACTGTCACCACTCTGCACTGC 3051  
421 GGCCAGACGTGTTAANCAAGGNTGCTGGCTGTGTTACTGTCACCACTCTGCACTGC 480  
3052 TGCTTTGAGACTCCATCCAGCCCGCAGGACGCCACCTGCTCCCTGAGCCTCCACTATCTCC 3111  
481 TGCTTTGAGACTCCANN 540  
3112 CTGTGACGGGTGAATCTCGTGTACTGTGTCTCGGGTCCATATATGAATTTGTGACAGGGT 3171  
541 CTGTGACGGGTGAATCTCGTGTACTGTGTCTCGGGTCCATATATGAATTTGTGACAGGGT 600  
3172 TCATCTATTTTAAACACAGATGTTTACAAATAAAGATTATTTCAAACCCCAAAAAA 3231  
601 TCATCTATTTTAAACACAGATGTTTACAAATAAAGATTATTTCAAACCCCAAAAAA 660  
3232 AAAAAAAAAAAAAA 3247  
661 AAAAAAAAAAAAAA 676

RESULT 13  
PATRAB15X 945 bp mRNA linear ROD 27-APR-1993  
DEFINITION Sprague-Dawley (clone LRB9), RAB15 mRNA, complete cds.  
ACCESSION M83679  
VERSION M83679.1 GI:206536  
KEYWORDS LMW GTP-binding protein.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 945)  
Elferink,L.A., Anzai,K. and Scheller,R.H.  
rab15, a novel low molecular weight GTP-binding protein  
specifically expressed in rat brain  
J. Biol. Chem. 267 (9), 5768-5775 (1992)  
92210533  
1313420  
PUBMED  
COMMENT Original source text: Rattus norvegicus (strain Sprague-Dawley)  
(library: LAMBDA ZAPII) adult brain cDNA to mRNA.  
FEATURES  
source  
1..945  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/tissue\_type="brain"  
/dev\_stage="adult"  
/tissue\_lib="LAMBDA ZAPII"  
1..945  
/gene="RAB15"  
220..858  
/gene="RAB15"  
/codon\_start=1  
/product="RAB15"  
/protein\_id="AAA41995.1"

/db\_xref="GI:206537"  
/translation="MAKQYDVLFRLLILGDSGVGKTKLLCRFTDNFHSHTIGVD  
FMKTIQVIRIQIWDTAGERYQITIKQYIRRAQCLIFLYDISSRSYQHMKW  
VSDVBYAPEGVKILIGNKADEBKROVREGOOLAKKEYGMDFVETSACTNLNKE  
SFTRLTELVLQARKELDGLRTCASNELALAELEDEGKTEGPANSSKTCWC"  
ORIGIN  
Query Match 17.7%; Score 577.6; DB 10; Length 945;  
Best Local Similarity 88.7%; Pred. No. 2.1e-133;  
Matches 637; Conservative 0; Mismatches 79; Indels 2; Gaps 1;  
QY 21 TTCCCGGCCCGCTGGCCCGATGTCGGAAGCAGTACGATGTGCTTCCGGCTGCTG 80  
Db 196 TTCCGGCCCGCTTGAAGTCCCTGCTGCGGAAACAGTACGATGTGCTTCCGGCTGCTG 255  
QY 81 CTGATCGGGGATCCCGGGTGGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140  
Db 256 CTGATCGGGGATCCCGGGTGGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 315  
QY 141 TTCCACTCTCGCACATCTCCACCATCGGTGTTGACTTTAAGATGAAGACCATAGAGGTA 200  
Db 316 TTCCACTCTCGCATATCTCCACCATCGGTGTTGACTTTAAGATGAAGACCATAGAGTA 375  
QY 201 GACGGCATCAAAGTGGGATACAGATCTGGGACACCTGCGGGGAGGAGGATACCGAGC 260  
Db 376 GACGGCATCAAAGTGGGATACAGATTTGGGACACAGCAGCGGGGAGGAGGATACCGAGC 435  
QY 261 ATCACAAGCAGTACTATCGGGGGCCCGGGGATATTTTGGTCTGATGACATTTAGCAGC 320  
Db 436 ATCACAAGCAGTACTATCGGGGGCCCGGGGATATTTTGGTCTGATGACATTTAGCAGC 495  
QY 321 GAGCGCTCTTACCAGCACATCATGAAAGTGGGTGAGTGCAGTGGATGAGTACGACACAGAA 380  
Db 496 GAGCGCTCTTATCAGCACATCATGAAAGTGGGTGAGTGCAGTGGATGAGTACGACACAGAA 555  
QY 381 GCGCTCAGAAAGATCTTATTGGGAATAAGCTGTGATGAGGAGCAGAAACCGGAGGTGGGA 440  
Db 556 GCGCTCAGAAAGATCTTATCGGGAATAAGCTGTGATGAGGAGCAGAAACCGGAGGTGGGG 615  
QY 441 AGAGACGAGGGCAGCAGCTGCGGAGGAGTATGGGATGGACTTCTATGAAAACAAGTGCC 500  
Db 616 AGAGACGAGGGCAGCAGCTGCGGATAGGATATGGGATGGACTTCTACGAAAACAAGTGCC 675  
QY 501 TGCACCAACCTCAACATTTAAAGAGTCACTCACCGCTGTGACAGAGCTGTGCTGCGAGCC 560  
Db 676 TGCACCAACCTTAACTTAAAGAGTCTTCACTCGTCTGACCGAGCTGTGCTGCGAGCT 735  
QY 561 CATAGGAAGAGCTGGAAGCCTCCGGATCGGTGCGGCAATGAGTTGGCACTGGCAGAG 620  
Db 736 CACAGGAAGAGCTGGATGGTCTCCGAACAATGTGCCAGCAATGAGCTCGCAGCTGGCCGAG 795  
QY 621 CTGGAGGAGGAGGAGGCAACCGGAGGCGCCAGCGAATCTTTCGAAAACCTGCTGCTGTC 680  
Db 796 CTGGAGGAACACGAGGCAAACTGAAAGGCCAGCAAACTCTTCAAGACCTGCTGCTGTC 855  
QY 681 TGA--GTCTGTGTGGGGCACCCACACGACACCCCTCTTCCCTCAGGAGCCCGTGG 736  
Db 856 TGAGGGTCTGTGTGGAAATCTCCCTCCCGAGTCCCTTCCCTCAGGAGCCCGTGG 913  
RESULT 14  
AX396088 481 bp DNA linear PAT 18-MAY-2002  
LOCUS AX396088  
DEFINITION Sequence 303 from Patent WO0212328.  
ACCESSION AX396088  
VERSION AX396088.1 GI:21066835  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS King,G.E., Meagher,M.J., Xu,J. and Secrist,H.

TITLE Compositions and methods for the therapy and diagnosis of colon cancer  
JOURNAL Patent: WO 0212328-A 303 14-FEB-2002;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers  
source 1..481  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 14.7%; Score 477.4; DB 6; Length 481;  
Best Local Similarity 99.4%; Pred. No. 2.6e-108;  
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 2582 GTTCTCTTGAAGATGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2641  
Db 1 GTTCTCTTGAAGATGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 60  
Qy 2642 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGGAAGTTA 2701  
Db 61 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGGAAGTTA 120  
Qy 2702 TATTGGGAGTGGTCCAAATCTGGGACCAAGTATTTCAAGCTTTCCCTGAAGATCAGGCG 2761  
Db 121 TATTGGGAGTGGTCCAAATCTGGGACCAAGTATTTCAAGCTTTCCCTGAAGATCAGGCG 180  
Qy 2762 GGTGCCATTCAATCTCTTCTCTCTAGCCCTCAGGAGGAGGACTATATTGTACT 2821  
Db 181 GGTGCCATTCAATCTCTTCTCTCTAGCCCTCAGGAGGAGGACTATATTGTACT 240  
Qy 2822 GTACCTAGGGGTTCTGGAAGGGAACAATGGAATCAGGATTTCTATAGACTGATAGGCC 2881  
Db 241 GTACCTAGGGGTTCTGGAAGGGAACAATGGAATCAGGATTTCTATAGACTGATAGGCC 300  
Qy 2882 TATCCACAAGGCCATGACTGGGAAAAGTATGGGAGCAGAAAGTGGGATTTTGG 2941  
Db 301 TATCCACAAGGCCATGACTGGGAAAAGTATGGGAGCAGAAAGTGGGATTTTGG 360  
Qy 2942 GGTGACGTACGCTCACCCTAACTTTTGGGGCTGGGCTCTTGGGAGGAGGACT 3001  
Db 361 GGTGACGTACGCTCACCCTAACTTTTGGGGCTGGGCTCTTGGGAGGAGGACT 420  
Qy 3002 GTTAAGCAGGCTCTGTGGCTGTTTACTTCGTACACCTCTGACACCTGCTGTGAGA 3061  
Db 421 GTTAACAGGCTCTGTGGCTGTTTACTTCGTACACCTCTGACACCTGCTGTGAGA 480  
Qy 3062 C 3062  
Db 481 C 481

RESULT 15  
AX334820/c  
LOCUS AX334820 463 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 5329 from Patent WO0194629.  
ACCESSION AX334820  
VERSION AX334820.1 GI:18125539  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrihan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 5329 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES Location/Qualifiers  
source 1..463  
/organism="Homo sapiens"

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 13.5%; Score 438.8; DB 6; Length 463;  
Best Local Similarity 99.3%; Pred. No. 1.2e-96;  
Matches 451; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
Qy 2790 CCCCTCAGGAAGAGGACTATATTGTACTGTACCTAGGGGTCTTGGAGGGAAC 2849  
Db 453 CCCCTCAGGAAGAGGACTATATTGTACTGTACCTAGGGGTCTTGGAGGGAAC 394  
Qy 2850 ATGAATCAGGATTTCTATAGACTATAGGCCCTATCCAAAGGCCATGACTGGGAAAG 2909  
Db 393 ATGAATCAGGATTTCTATAGACTATAGGCCCTATCCAAAGGCCATGACTGGGAAAG 334  
Qy 2910 GTATGGGAGCAGAGGAGAAATTTGGGATTTAGGGTGCAGCTACCTCACCTAACTTTT 2969  
Db 333 GTATGGGAGCAGAGGAGAAATTTGGGATTTAGGGTGCAG-TACGCTCACCTAACTTTT 275  
Qy 2970 GGTGGCTTGGGGCATGTCTTGGGCCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTAC 3029  
Db 274 GGTGGCTTGGGGTATGTCTTGGGCCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTAC 215  
Qy 3030 TCGTCACCACTCTGCACCTGCTGTCTTGGAGACTCCATCCAGCCCCAGGCCACCTG 3089  
Db 214 TCGTCACCACTCTGCACCTGCTGTCTTGGAGACTCCATCCAGCCCCAGGCCACCTG 155  
Qy 3090 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTGTGTACTGTGTCTGGGTCC 3149  
Db 154 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTGTGTACTGTGTCTGGGTCC 95  
Qy 3150 ATATATGAATTTGTGAGCAGGCTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGAT 3209  
Db 94 ATATATGAATTTGTGAGCAGGCTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGAT 35  
Qy 3210 TATTTCAAACCCACCAAAAAAATAAAAAA 3243  
Db 34 TATTTCAAACCCACCAAAAAAATAAAAAA 1

Search completed: March 21, 2004, 08:59:33  
Job time : 12136 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 02:17:36 ; Search time 1147 Seconds  
(without alignments)  
12063.105 Million cell updates/sec

Title: US-09-817-198C-1

Perfect score: 3257

Sequence: 1 tgccgctgcgcgcgcgcag.....aaaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: geneseqn1980s:\*

2: geneseqn1990s:\*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001bs:\*

6: geneseqn2002s:\*

7: geneseqn2003as:\*

8: geneseqn2003bs:\*

9: geneseqn2003cs:\*

10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3257	100.0	3257	7 AAD47168	Aad47168 Human Ras
2	2696.6	82.8	3240	4 AAK82285	Aak82285 Human imm
3	2696.6	82.8	28770	7 AAD47169	Aad47169 Human Ras
4	1568.2	48.1	1666	3 AAC75813	Aac75813 Human ORF
5	1192.8	36.6	7924	6 ABK40069	Abk40069 Human che
6	1192.8	36.6	7924	6 ABK4138	Abk4138 Human imm
7	1160.8	35.6	7924	6 ABK40070	Abk40070 Human che
8	1160.8	35.6	7924	6 ABK4139	Abk4139 Human imm
9	1160.6	35.6	2021	4 AAS27053	Aas27053 CDNA enco
10	1160.6	35.6	2021	9 ADB93231	Adb93231 Human cDN
11	1000.2	30.7	1054	6 ABQ93361	Abq93361 Human cDN
12	866.2	26.6	895	2 AAZ24400	Aaz24400 Human bla
13	534.2	16.4	566	4 AAS27472	Aas27472 CDNA enco
14	534.2	16.4	566	9 ADB93650	Adb93650 Human cDN
15	477.4	14.7	481	6 ABK44752	Abk44752 CDNA enco
16	438.8	13.5	463	6 ABL66992	Abi66992 Thyroid c
17	424.4	13.0	458	8 ACH33636	Ach33636 Human end
18	313.4	9.6	320	2 AAT25457	Aat25457 Human gen
19	307	9.4	313	6 ABL38341	Abi38341 Human col
20	229	7.0	1986	4 ABA09160	Abao9160 Human rab
21	228.4	7.0	2411	6 ABQ54410	Abq54410 Human ova
22	217.8	6.7	1161	4 AAH75182	Aah75182 Nucleotid
23	217.8	6.7	1265	6 ABV78060	Abv78060 Hypoxia-r

24	217.8	6.7	2497	4	AAS60878	Human can
25	217.8	6.7	2497	4	AAS60884	Human can
26	217.8	6.7	2497	4	AAS60893	Human can
27	217.8	6.7	2497	4	AAS60895	Human can
28	217.8	6.7	2497	5	ABV25781	Human pro
29	217.8	6.7	2497	5	ABV30037	Human pro
30	217.8	6.7	3077	4	AAH13912	Human cDN
31	214.2	6.6	866	4	AAH04301	Human cDN
32	209	6.4	1859	8	ADB12721	Human pol
33	203.2	6.2	991	9	ADB52880	Primary r
34	202.6	6.2	412	3	AAC01700	Human sec
35	200.2	6.1	716	3	AAA40104	Human Rab
36	200.2	6.1	861	3	AAA40108	Human Rab
37	198.4	6.1	956	3	AAAG6887	Nucleotid
38	198.4	6.1	1537	3	AAF16196	Human pro
39	198.4	6.1	3533	4	AAH17889	Human cDN
40	196.8	6.0	674	5	AAST1453	Human cDN
41	196.6	6.0	2247	4	ABL29661	Drosophil
42	195.6	6.0	3164	6	ABX04179	Human mRN
43	195.2	6.0	888	3	AAA40109	Canine Ra
44	192	5.9	1274	3	AAF18072	Lung canc
45	188.6	5.8	893	5	AAAS7552	DNA enco

ALIGNMENTS

RESULT 1  
RAD47168  
ID AAD47168 standard; cDNA; 3257 BP.  
XX  
AC AAD47168;  
XX  
DT 24-FEB-2003 (first entry)  
XX  
DE Human Ras-like protein encoding cDNA.

Human; Ras-like protein; inflammation; cell proliferation; apoptosis;  
immunodeficiency; neurodegenerative; Alzheimer's disease; cirrhosis;  
Parkinson's disease; wasting disease; cachexia; myocardial infarction;  
osteoporosis; atherosclerosis; glomerulonephritis; Crohn's disease;  
irritable bowel syndrome; multiple sclerosis; osteoarthritis; allergy;  
pancreatitis; autoimmune thyroiditis; rheumatoid arthritis; osteopathic;  
Sjogren's syndrome; infection; transgenic; gene therapy; neotropic; gout;  
neuroprotective; vasotonic; cytostatic; dermatological; nephrotropic;  
ophthalmological; tranquiliser; cancer; stroke; Grave's disease; AIDS;  
asthma; anaemia; drug screening; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH 5'UTR 1..44  
FT CDS /\*tag= a  
FT CDS 45..683  
FT /\*tag= b  
FT 3'UTR /product= "Human Ras-like protein"  
684..3257  
/\*tag= c

XX WO200277193-A2.

XX 03-OCT-2002.

XX 27-MAR-2002; 2002WO-US009328.

XX 27-MAR-2001; 2001US-00817198.

XX (PEKE ) PE CORP.

XX Gan W, Ye J, Di Francesco V, Beasley EM;

XX WPI; 2003-018913/01.

XX P-PSDB; AAE29096.

XX New isolated human Ras-like protein polypeptide, useful for diagnosing,  
PT treating or preventing inflammation and disorders associated with cell  
PT proliferation and apoptosis, e.g. AIDS, cancer, allergies, anemia, asthma  
PT or stroke.

XX Claim 4; Page 73-74; 82pp; English.

XX The invention relates to human Ras-like protein and its corresponding  
CC nucleic acid. The Ras-like protein and DNA is useful in the development  
CC of human therapeutics and diagnostic compositions. They are useful in the  
CC diagnosis, prevention and treatment of inflammation and disorders  
CC associated with cell proliferation and apoptosis, e.g. AIDS and other  
CC infectious or genetic immunodeficiencies, neurodegenerative disease e.g.  
CC Alzheimer's disease, Parkinson's disease, wasting diseases e.g. cachexia,  
CC ischaemic injuries e.g. myocardial infarction, stroke or reperfusion  
CC injury, toxin-induced diseases such as alcohol-induced liver damage or  
CC cirrhosis, osteoporosis or cancer. They are also used to treat disorders  
CC associated with inflammation including allergies, atopic dermatitis,  
CC atherosclerosis, asthma, anaemia, Crohn's disease, diabetes mellitus,  
CC Grave's disease, glomerulonephritis, gout, irritable bowel syndrome,  
CC lupus erythematosus, multiple sclerosis, osteoarthritis, pancreatitis,  
CC autoimmune thyroiditis, rheumatoid arthritis, Sjogren's syndrome,  
CC uveitis, trauma, or viral, bacterial, fungal, parasitic, protozoal or  
CC helminthic infections. The antibodies of the invention are useful in  
CC pharmacogenomic analysis or for tissue typing. The transgenic animals are  
CC useful for studying the function of a Ras-like protein, and identifying  
CC and evaluating modulators of its activity. Ras-like protein is used in  
CC drug screening assays and its DNA is used in gene therapy. The present  
CC sequence is human Ras-like protein encoding cDNA

XX Sequence 3257 BP; 793 A; 892 C; 870 G; 702 T; 0 U; 0 Other;

Query Match 100.0%; Score 3257; DB 7; Length 3257;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TGCCCGCTGCGCCGCGCAGTTCCCGCCCGCGTGGCCCGCAGTCATCGCGGAAGCAGTACG	60
DB	1	TGCCCGCTGCGCCGCGCAGTTCCCGCCCGCGTGGCCCGCAGTCATCGCGGAAGCAGTACG	60
QY	61	ATGTCGTGTTCCGCTGCTGATCGGGGATCCCGGGGTGGGCAAGACCTGCTGCTGT	120
DB	61	ATGTCGTGTTCCGCTGCTGATCGGGGATCCCGGGGTGGGCAAGACCTGCTGCTGT	120
QY	121	GCCTGTTACCGCAAGAGTTCCACTCTCTCGCATCTCCACCATCGGTGTCACCTTTA	180
DB	121	GCCTGTTACCGCAAGAGTTCCACTCTCTCGCATCTCCACCATCGGTGTTGACTTTA	180
QY	181	AGATGAAGACCATAGAGTACGCGCATCAAGTGGGATACAGATCTGGGACATCTGCAG	240
DB	181	AGATGAAGACCATAGAGTACGCGCATCAAGTGGGATACAGATCTGGGACATCTGCAG	240
QY	241	GGCAGGAGATACCAAGCATCAAGCAGTACTATCGCGGGGCCAGGGGATATTTT	300
DB	241	GGCAGGAGATACCAAGCATCAAGCAGTACTATCGCGGGGCCAGGGGATATTTT	300
QY	301	TGGTCTATGATACATAGCAGGCGCTTACAGGACATCATGAAGTGGTCACTGACG	360
DB	301	TGGTCTATGATACATAGCAGGCGCTTACAGGACATCATGAAGTGGTCACTGACG	360
QY	361	TGGATGATACGACCAAGAGGCGTCCAGGAAGATCTTATTTGGGAATAAGCTGATGAG	420
DB	361	TGGATGATACGACCAAGAGGCGTCCAGGAAGATCTTATTTGGGAATAAGCTGATGAG	420
QY	421	AGCAAAACGCGAGTGGGAGAGCAAGGCGAGCAGCTGGGCGAGGATATGGCATGG	480
DB	421	AGCAAAACGCGAGTGGGAGAGCAAGGCGAGCAGCTGGGCGAGGATATGGCATGG	480
QY	481	ACTTCTATGAACAAGTGGCTGACCAACCTTACATTTAAGAGTCACTTCAACCGCTGA	540
DB	481	ACTTCTATGAACAAGTGGCTGACCAACCTTACATTTAAGAGTCACTTCAACCGCTGA	540

QY	541	CAGAGCTGCTGCTCGAGGCCATAGGAAGAGCTGGAAGGCTCCGATGCTGCCAGCA	600
DB	541	CAGAGCTGCTGCTCGAGGCCATAGGAAGAGCTGGAAGGCTCCGATGCTGCCAGCA	600
QY	601	ATGAGTTGGCACTGGCGAGAGCTGGAGAGGAGGCAAAACCCGAGGGGCCAGCAACT	660
DB	601	ATGAGTTGGCACTGGCGAGAGCTGGAGAGGAGGCAAAACCCGAGGGGCCAGCAACT	660
QY	661	CTTCGAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720
DB	661	CTTCGAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720
QY	721	CTCAGGAGGCGCTGGGCGACAGAGGAGGCGGGGCTTTCGCTGCTGCTGCTGCTGCT	780
DB	721	CTCAGGAGGCGCTGGGCGACAGAGGAGGCGGGGCTTTCGCTGCTGCTGCTGCTGCT	780
QY	781	GTGATGACCTATTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	840
DB	781	GTGATGACCTATTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	840
QY	841	CTGCTGCTCATCTCAAGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900
DB	841	CTGCTGCTCATCTCAAGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	900
QY	901	GCCTGTTTCCCGAGCCACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960
DB	901	GCCTGTTTCCCGAGCCACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960
QY	961	ACCATCCCGCACCCACAGAGGCTGAGTCCAGGCTGAGTCCAGGCTGAGTCCAGGCTGAGT	1020
DB	961	ACCATCCCGCACCCACAGAGGCTGAGTCCAGGCTGAGTCCAGGCTGAGTCCAGGCTGAGT	1020
QY	1021	CTTTCTCCGTCATCGTCTGCT	1080
DB	1021	CTTTCTCCGTCATCGTCTGCT	1080
QY	1081	TGACCCCTCCCTCCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
DB	1081	TGACCCCTCCCTCCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
QY	1141	CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1200
DB	1141	CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1200
QY	1201	GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260
DB	1201	GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260
QY	1261	GGCCAAAGGCTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1320
DB	1261	GGCCAAAGGCTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1320
QY	1321	TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1380
DB	1321	TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1380
QY	1381	CCGGGGAGTCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
DB	1381	CCGGGGAGTCTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
QY	1441	CCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1500
DB	1441	CCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1500
QY	1501	AGAAGGAAACAAATAGAGGGGTGGCGGATATAAAGTCACTTCTTCTTCTTCTTCTTCTTCT	1560
DB	1501	AGAAGGAAACAAATAGAGGGGTGGCGGATATAAAGTCACTTCTTCTTCTTCTTCTTCTTCT	1560
QY	1561	TGCAGCATGAACAATTTCTCTCCACCTGGCTCCCAATTTTAAAGATGAGGACCAAGGC	1620
DB	1561	TGCAGCATGAACAATTTCTCTCCACCTGGCTCCCAATTTTAAAGATGAGGACCAAGGC	1620
QY	1621	CTGTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCTAGGCCAACCAT	1680

Db 1621 |||||CTGTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGGTCAGTGACACTGTGACGGCCAAACCAT 1680  
Qy 1681 GCACCTCCACAAAGGGAGACTTTTGGAAATGAAGACTAGCTCCTATGTATCAGGTAAAGA 1740  
Db 1681 GCACCTCCACAAAGGGAGACTTTTGGAAATGAAGACTAGCTCCTATGTATCAGGTAAAGA 1740  
Qy 1741 GCAAGGAGAGCTGGCCAGGCAAGCAGCTTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800  
Db 1741 GCAAGGAGAGCTGGCCAGGCAAGCAGCTTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800  
Qy 1801 GGGCTCTTAGGCCCCCATCTTCCATTTCTTAGTTAAGAGAGCATTTTCTCAGACTGCCA 1860  
Db 1801 GGGCTCTTAGGCCCCCATCTTCCATTTCTTAGTTAAGAGAGCATTTTCTCAGACTGCCA 1860  
Qy 1861 GGGGAGGACTAGCCTAGCCTTCCAGCAACCAAGGTTCTCTGGGACCCCAAAGTTTATGG 1920  
Db 1861 GGGGAGGACTAGCCTAGCCTTCCAGCAACCAAGGTTCTCTGGGACCCCAAAGTTTATGG 1920  
Qy 1921 GAGAAGGGCAAGACTTTTCATGGGAGAGAGGAAGGCCCTGGGTAGAAACGCTTGGTG 1980  
Db 1921 GAGAAGGGCAAGACTTTTCATGGGAGAGAGGAAGGCCCTGGGTAGAAACGCTTGGTG 1980  
Qy 1981 CTGTTCTTTGGGCTTTAAGCAAAAGCGCTCATCTTGGCCCTTACCTCTCTGATAGGCTT 2040  
Db 1981 CTGTTCTTTGGGCTTTAAGCAAAAGCGCTCATCTTGGCCCTTACCTCTCTGATAGGCTT 2040  
Qy 2041 GAGGTTTGGCCAAACCACTGTGGCTACAGGTGGAGGGAAGAGACTCTCTTCCCTCAGAG 2100  
Db 2041 GAGGTTTGGCCAAACCACTGTGGCTACAGGTGGAGGGAAGAGACTCTCTTCCCTCAGAG 2100  
Qy 2101 TGCATATGTTTCAGGAAGTTTCTTTAAACCCCATATGGCCCAAGAGTAGCTCGTAGGAGGCC 2160  
Db 2101 TGCATATGTTTCAGGAAGTTTCTTTAAACCCCATATGGCCCAAGAGTAGCTCGTAGGAGGCC 2160  
Qy 2161 TTTAAAGACGGAACAAGTAATTTTACCAGTTTCTTACAGTTTCTTACAGTTTCTTACAGTTT 2220  
Db 2161 TTTAAAGACGGAACAAGTAATTTTACCAGTTTCTTACAGTTTCTTACAGTTTCTTACAGTTT 2220  
Qy 2221 TGGCGAGGCTTAGGAAGAGGTCATCTTAAAGCCACACATTTAGCTGCACCTGCGTGGCTG 2280  
Db 2221 TGGCGAGGCTTAGGAAGAGGTCATCTTAAAGCCACACATTTAGCTGCACCTGCGTGGCTG 2280  
Qy 2281 CAGCCAAAACAAAGAACTGGGTGTAGTATTCATCAACTAAGAACCAAAATCCAGGGCA 2340  
Db 2281 CAGCCAAAACAAAGAACTGGGTGTAGTATTCATCAACTAAGAACCAAAATCCAGGGCA 2340  
Qy 2341 CTCATATGTGAAGATAGAAACCTCCTTCTTACTCTCTCAAAAAGAGTGGGGAAGA 2400  
Db 2341 CTCATATGTGAAGATAGAAACCTCCTTCTTACTCTCTCAAAAAGAGTGGGGAAGA 2400  
Qy 2401 ACCATCAAACTTTCTCTCTGACTTACCAAAACAGGAAAACAGCAGGAGGGTGGCTCA 2460  
Db 2401 ACCATCAAACTTTCTCTCTGACTTACCAAAACAGGAAAACAGCAGGAGGGTGGCTCA 2460  
Qy 2461 GGACTTAGGGACAGGTATAGCTTAGTGTGGAAAGCAAGAGGAGAGCAGGAAGTTGTAA 2520  
Db 2461 GGACTTAGGGACAGGTATAGCTTAGTGTGGAAAGCAAGAGGAGAGCAGGAAGTTGTAA 2520  
Qy 2521 ATCACTGGCTAATGAGAAAAGGAGACAGCTAACTTAGGTGAAGCTGTGACTAGGCTGG 2580  
Db 2521 ATCACTGGCTAATGAGAAAAGGAGACAGCTAACTTAGGTGAAGCTGTGACTAGGCTGG 2580  
Qy 2581 AGTTGCTTCTTGAAGTGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640  
Db 2581 AGTTGCTTCTTGAAGTGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640  
Qy 2641 GCTAGGGAAGTAGGTGATGCGAGCCCTCAAGTCTGTCTTTCAGCCAGGAGCTTGAAGATT 2700  
Db 2641 GCTAGGGAAGTAGGTGATGCGAGCCCTCAAGTCTGTCTTTCAGCCAGGAGCTTGAAGATT 2700  
Qy 2701 ATATTGGGCACTGGCTCAATCTGTGGACCAAGTATTTCCCTGAAAGATCAGGCA 2760  
|||||

Db 2701 ATATTGGGCACTGGCTCAATCTGTGGACCAAGTATTTTCCCTTCCCTGAAAGATCAGGCA 2760  
Qy 2761 GGGTGCCATTCAATTGTCTTCTCTCTAGCCCCCTCAGGAAAAGAGCACTATATTGTATC 2820  
Db 2761 GGGTGCCATTCAATTGTCTTCTCTCTAGCCCCCTCAGGAAAAGAGCACTATATTGTATC 2820  
Qy 2821 TGTTACCTTAGGGTTCTGGAAAGGAAAACATGGAAATCAGGAATCTATAGACTGATAGGCC 2880  
Db 2821 TGTTACCTTAGGGTTCTGGAAAGGAAAACATGGAAATCAGGAATCTATAGACTGATAGGCC 2880  
Qy 2881 CTATCCACAAAGGGCCATGACTTGGGAAAGGATATGGGAGCAGAGGAATTTGGGATTTTA 2940  
Db 2881 CTATCCACAAAGGGCCATGACTTGGGAAAGGATATGGGAGCAGAGGAATTTGGGATTTTA 2940  
Qy 2941 GGGTGCACTACGCTCAACCCCTAAACTTTTGTGGCCCTGGGGCATGTCTTGAAGGCCAGAC 3000  
Db 2941 GGGTGCACTACGCTCAACCCCTAAACTTTTGTGGCCCTGGGGCATGTCTTGAAGGCCAGAC 3000  
Qy 3001 TGTTAAGCAGGCTCTGTGGCTGTGCTTCTGCTTCAACCTCTCTGCACTGTCTGTCTTGAG 3060  
Db 3001 TGTTAAGCAGGCTCTGTGGCTGTGCTTCTGCTTCAACCTCTCTGCACTGTCTGTCTTGAG 3060  
Qy 3061 ACTCCATCCAGCCCCAGGCAAGCCACCTGCTCCTGAGCCTCCACTATCTCCTGTGAGGG 3120  
Db 3061 ACTCCATCCAGCCCCAGGCAAGCCACCTGCTCCTGAGCCTCCACTATCTCCTGTGAGGG 3120  
Qy 3121 GTGAACCTTCTGTACTGTGTCTCGGGTCCATATATGAATTTGTGAGCAGGGTTCACTATT 3180  
Db 3121 GTGAACCTTCTGTACTGTGTCTCGGGTCCATATATGAATTTGTGAGCAGGGTTCACTATT 3180  
Qy 3181 TTAACACACAGATGTTTACAAAATAAAGATTTATTTCAAACCAACCAAAAAAAAAAAAA 3240  
Db 3181 TTAACACACAGATGTTTACAAAATAAAGATTTATTTCAAACCAACCAAAAAAAAAAAAA 3240  
Qy 3241 AAAAAAAAAAAAAAAAAA 3257  
Db 3241 AAAAAAAAAAAAAAAAAA 3257

RESULT 2  
AAK82285/c  
ID AAK82285 standard; DNA; 3240 BP.  
XX  
AC AAK82285;  
XX  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37097.  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
FN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205151P.  
PR 27-JUN-2000; 2000US-0209457P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.



CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX

SQ Sequence 3240 BP; 713 A; 882 C; 887 G; 758 T; 0 U; 0 Other;  
Query Match 82.8%; Score 2696.6; DB 4; Length 3240;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 271; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy	473	TGGCATGGACTTCTATGAACAAGTGGCTGCACCAACCTCAACATTAAAGATCAATTCAC	532
Db	3100	TGCCAGGGAGAGGAGGAGACACTGGACTAACCTGTGCCCTTTTGGTTTCCAGTCAATTCAC	3041
Qy	533	CGCTCTGACAGAGCTGTGTCTGAGGCCCATAGGAAGAGCTGAAGCCCTCCGGATGCG	592
Db	3040	CGCTCTGACAGAGCTGTGTCTGAGGCCCATAGGAAGAGCTGAAGCCCTCCGGATGCG	2981
Qy	593	TGCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGAGGAGGCAAAACCCGAGGGCCC	652
Db	2980	TGCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGAGGAGGCAAAACCCGAGGGCCC	2921
Qy	653	AGCGAACTCTTCGAAACCTGCTGTGCTGAGTCTCTGTGTGGGCGACCCCAACAGACACC	712
Db	2920	AGCGAACTCTTCGAAACCTGCTGTGCTGAGTCTCTGTGTGGGCGACCCCAACAGACACC	2861
Qy	713	CCTCTTCCCTCAGAGGCCGTGGGCGACAGAGGAGCGGGGCTTTGGCCCTGTCTGTGT	772
Db	2860	CCTCTTCCCTCAGAGGCCGTGGGCGACAGAGGAGCGGGGCTTTGGCCCTGTCTGTGT	2801
Qy	773	CCTCTCTGTGATGACACCTATTGAGTATCAGTAGCCACTACCCCTGCTGGCCCTGA	832
Db	2800	CCTCTCTGTGATGACACCTATTGAGTATCAGTAGCCACTACCCCTGCTGGCCCTGA	2741
Qy	833	GAGGGGCTCTGCTGTCTATCTCAAGCAGCCCTGTCTCCAGCCCTCCACCTCGAGTGGT	892
Db	2740	GAGGGGCTCTGCTGTCTATCTCAAGCAGCCCTGTCTCCAGCCCTCCACCTCGAGTGGT	2681
Qy	893	CTTCTTCAGCCCTGTTTCCAGCCACACAGGCTGTACGACCCCAACGATGCGCGAAGC	952
Db	2680	CTTCTTCAGCCCTGTTTCCAGCCACACAGGCTGTACGACCCCAACGATGCGCGAAGC	2621
Qy	953	ACTGTCTCACCATCCGACCCACAGACACAGCAGGCGTGGAGTCCAGGCCACTTTC	1012
Db	2620	ACTGTCTCACCATCCGACCCACAGACACAGCAGGCGTGGAGTCCAGGCCACTTTC	2561
Qy	1013	AGCTGTCTCTTCTCCGCTGATCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1072
Db	2560	AGCTGTCTCTTCTCCGCTGATCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2501
Qy	1073	TCCTTCTCTGACCCCTCCCTCCGCTGCTTCTGATCAAAAGCTCTCAAAACCCCGTCCC	1132
Db	2500	TCCTTCTCTGACCCCTCCCTCCGCTGCTTCTGATCAAAAGCTCTCAAAACCCCGTCCC	2441
Qy	1133	CCGTGTCTCTGCTGTGTGAGCTCGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1192
Db	2440	CCGTGTCTCTGCTGTGTGAGCTCGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2381
Qy	1193	GGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCCACCCCTGCT	1252
Db	2380	GGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCCACCCCTGCT	2321
Qy	1253	CGTGGTGGGCGAAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1312
Db	2320	CGTGGTGGGCGAAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2261
Qy	1313	ATGTGCCATGGGCTGTCTCTCCCAAGTACCTGTGCAAGTGTGAGCATCGAGGTAGAGGGA	1372
Db	2260	ATGTGCCATGGGCTGTCTCTCCCAAGTACCTGTGCAAGTGTGAGCATCGAGGTAGAGGGA	2201

Qy	1373	AACAGCAACCGGGGAGTCTCTGAGCCTGGGGCTGCCCTACCTCTACCCATTTCCCGACCA	1432
Db	2200	AACGGCAACCGGGGAGTCTCTGAGCCTGGGGCTGCCCTACCTCTACCCATTTCCCGACCA	2141
Qy	1433	GAGCTTTGCCCTTGTCTTGGCTGCCCGCTGCTCTTTTGGGAACTGAGCTCAGAGGAGG	1492
Db	2140	GAGCTTTGCCCTTGTCTTGGCTGCCCGCTGCTCTTTTGGGAACTGAGCTCAGAGGAGG	2081
Qy	1493	TGCTTTCAGAGAGGAAACAAAATGAGGGGTGGCAGGAGTAAAAAGTCACTCTTCTCT	1552
Db	2080	TGCTTTCAGAGAGGAAACAAAATGAGGGGTGGCAGGAGTAAAAAGTCACTCTTCTCT	2021
Qy	1553	ACTTCCCATGAGCATGAACACAAATTTCTCTCACTGGCTCCCAATTTAAAGATGCG	1612
Db	2020	ACTTCCCATGAGCATGAACACAAATTTCTCTCACTGGCTCCCAATTTAAAGATGCG	1961
Qy	1613	ACCAAGCCCTGTGGGTACTCCAGGGGCAAGAGAGCCCTGGGTCAGTGACACTGTGAGG	1672
Db	1960	ACCAAGCCCTGTGGGTACTCCAGGGGCAAGAGAGCCCTGGGTCAGTGACACTGTGAGG	1901
Qy	1673	CCAACCATGCACTCCACAAAGGGGAGCAATTTGGAATGAAGGACTAGCTCTATGTATCA	1732
Db	1900	CCAACCATGCACTCCACAAAGGGGAGCAATTTGGAATGAAGGACTAGCTCTATGTATCA	1841
Qy	1733	GGTTAAGAGCAAGGAGAGCTGGCCAGGACAGCAGTTTGACAGAGAGGGAATGTAG	1792
Db	1840	GGTTAAGAGCAAGGAGAGCTGGCCAGGACAGCAGTTTGACAGAGAGGGAATGTAG	1781
Qy	1793	CAACAGAGAGCCCTCTTAGGCCCATCTTCCATTTCTTAGGTAAGAAGAGCATTTCTCTCA	1852
Db	1780	CAACAGAGAGCCCTCTTAGGCCCATCTTCCATTTCTTAGGTAAGAAGAGCATTTCTCTCA	1721
Qy	1853	GACTTCCAGCGGAGGACTGAGCCTAGCCTTTCAGCAACCAAGGTTCTCTGGGACCCAAA	1912
Db	1720	GACTTCCAGCGGAGGACTGAGCCTAGCCTTTCAGCAACCAAGGTTCTCTGGGACCCAAA	1661
Qy	1913	GTTTATGGGAGAGGGCAAGAGACTTCAATGGGAAGAGAGGAGGCTCGGTAGAGAAC	1972
Db	1660	GTTTATGGGAGAGGGCAAGAGACTTCAATGGGAAGAGAGGAGGCTCGGTAGAGAAC	1601
Qy	1973	GCTTGTGTCTGTCTCTTTGGCCCTTTAAGACAAAGCGCTCATCTTGCCTCTACCTCTG	2032
Db	1600	GCTTGTGTCTGTCTCTTTGGCCCTTTAAGACAAAGCGCTCATCTTGCCTCTACCTCTG	1541
Qy	2033	ATAGCTTTGAGGGTTTCCCAACACACTGTGGCTTACAGGTGGAGGAGAGAGCTCTCTTC	2092
Db	1540	ATAGCTTTGAGGGTTTCCCAACACACTGTGGCTTACAGGTGGAGGAGAGAGCTCTCTTC	1481
Qy	2093	CTCCAGAGTGTATGTTTCAAGAGTCTTTTAAACCCCATATGGCCCAAGAGTAGCTCGTA	2152
Db	1480	CTCCAGAGTGTATGTTTCAAGAGTCTTTTAAACCCCATATGGCCCAAGAGTAGCTCGTA	1421
Qy	2153	GGAGGCCCTTTAAGAGCGGAACAAAGTAATTTAACAAGTTCTAGTGGGTCTCTGCCACCG	2212
Db	1420	GGAGGCCCTTTAAGAGCGGAACAAAGTAATTTAACAAGTTCTAGTGGGTCTCTGCCACCG	1361
Qy	2213	TCCCAAGCTGGCGAGCGCTTAGGAAGAGGGTCAATTTAAAGCCACATTAAGTGCACCTG	2272
Db	1360	TCCCAAGCTGGCGAGCGCTTAGGAAGAGGGTCAATTTAAAGCCACATTAAGTGCACCTG	1301
Qy	2273	CGTGGCTGACGACCAAAACAAAGAACTGGGTGTTGAGTATTCATCAACTAAGAACCAAAAT	2332
Db	1300	CGTGGCTGACGACCAAAACAAAGAACTGGGTGTTGAGTATTCATCAACTAAGAACCAAAAT	1241
Qy	2333	CCAGGCACTCATATGTGAAGGATAAGAACTCTCACTTCTCTTCTCTCTCTCTCTCTCTCTCT	2392
Db	1240	CCAGGCACTCATATGTGAAGGATAAGAACTCTCACTTCTCTTCTCTCTCTCTCTCTCTCTCT	1181
Qy	2393	GGGAAAGAACCATCAAAACCTTCTCTGACTTACCAACACAGGAAACACAGAGGAGG	2452
Db	1180	GGGAAAGAACCATCAAAACCTTCTCTGACTTACCAACACAGGAAACACAGAGGAGG	1121
Qy	2453	GTGGCTCAGGACTTAGGGACAGGGGTATAGCTTAGATGGTGAAAGCAAGGAGGAGGAGGA	2512

Db 1120 GTGGCTCAGGACTTAGGGACAGGGGTATAGCTTAGATGGTGAAGCAAGGAGCAGGGA 1061  
Qy 2513 AGTTGTAATCACTGGCTTAATGAAAGAGAGACAGCTAACTTAGATGAAGCTGTGAC 2572  
Db 1060 AGTTGTAATCACTGGCTTAATGAAAGAGAGACAGCTAACTTAGATGAAGCTGTGAC 1001  
Qy 2573 TAGGCTGGAGTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 2632  
Db 1000 TAGGCTGGAGTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 941  
Qy 2633 ACATGGGCTAGGAAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGACTT 2692  
Db 940 ACATGGGCTAGGAAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGACTT 881  
Qy 2693 GAGAAAGTATATGGGCAGTGGCTCCATCTGTGGACAGTATTTACAGCTTCCCTGAAG 2752  
Db 880 GAGAAAGTATATGGGCAGTGGCTCCATCTGTGGACAGTATTTACAGCTTCCCTGAAG 821  
Qy 2753 ATCAGGCAGGCTGCATTTGCTTCTTCTCTAGCCCTCAGGAAAGAGGACTAT 2812  
Db 820 ATCAGGCAGGCTGCATTTGCTTCTTCTCTAGCCCTCAGGAAAGAGGACTAT 761  
Qy 2813 ATTGTACTGTACCTTAGGGGTTCTGGAAGGGAAGAAACATGGAATTCATAGACT 2872  
Db 760 ATTGTACTGTACCTTAGGGGTTCTGGAAGGGAAGAAACATGGAATTCATAGACT 701  
Qy 2873 GATAGGCCCTATCCAGAGGCCATGCTGGGAAAGGTATGGAGCAGAGGAATTG 2932  
Db 700 GATAGGCCCTATCCAGAGGCCATGCTGGGAAAGGTATGGAGCAGAGGAATTG 641  
Qy 2933 GGATTTTAGGCTGAGCTACGCTCACCTAACTTTTGGTGGCTGGGGCATGCTTTGAG 2992  
Db 640 GGATTTTAGGCTGAGCTACGCTCACCTAACTTTTGGTGGCTGGGGCATGCTTTGAG 581  
Qy 2993 GCCCAGACTGTAAAGCAGGCTCTGCTGGCTGTTTACTGTGCTACCACTCTGACCTGCT 3052  
Db 580 GCCCAGACTGTAAAGCAGGCTCTGCTGGCTGTTTACTGTGCTACCACTCTGACCTGCT 521  
Qy 3053 GTCTTGAGACTCCATCCAGCCAGCCAGCCAGCTCTCTGAGCTCCCATCTATCTCCC 3112  
Db 520 GTCTTGAGACTCCATCCAGCCAGCCAGCCAGCTCTCTGAGCTCCCATCTATCTCCC 461  
Qy 3113 TGTGACGGGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3172  
Db 460 TGTGACGGGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401  
Qy 3173 CATCTATTTTAAACACAGATGTTTACAAAATAAAGATTTTCAAAACCACC 3223  
Db 400 CATCTATTTTAAACACAGATGTTTACAAAATAAAGATTTTCAAAACCACC 350

RESULT 3

AAD47169  
ID AAD47169 standard; DNA; 28770 BP.

AC AAD47169;

XX 24-FEB-2003 (first entry)

DE Human Ras-like protein encoding gene.

XX Human; Ras-like protein; inflammation; cell proliferation; apoptosis;  
XX immunodeficiency; neurodegenerative; Alzheimer's disease; cirrhosis;  
XX Parkinson's disease; wasting disease; cachexia; myocardial infarction;  
XX osteoporosis; atherosclerosis; glomerulonephritis; Crohn's disease;  
XX irritable bowel syndrome; multiple sclerosis; osteoarthritis; allergy;  
XX pancreatitis; autoimmune thyroiditis; rheumatoid arthritis; osteopathic;  
XX Sjogren's syndrome; infection; transgenic; gene therapy; neotropic; gout;  
XX neuroprotective; vasotropic; cytotostatic; dermatological; nephrotropic;  
XX ophthalmological; tranquilliser; cancer; stroke; Grave's disease; AIDS;  
XX asthma; anaemia; drug screening; gene; chromosome 14; ds.

OS Homo sapiens.  
XX Key Location/Qualifiers  
FH variation replace(206, -)  
FT /\*tag= a  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT 2044..25772  
FT /\*tag= b  
FT /\*product= "Human Ras-like protein"  
FT 2044..2167  
FT /\*tag= c  
FT intron 2168..21554  
FT /\*tag= d  
FT variation replace(4963,T)  
FT /\*tag= e  
FT variation replace(8175,A)  
FT /\*tag= f  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(10515,C)  
FT /\*tag= g  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(13034,C)  
FT /\*tag= h  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(13781,C)  
FT /\*tag= i  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(14050,C)  
FT /\*tag= j  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(14273, -)  
FT /\*tag= k  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(17582,C)  
FT /\*tag= l  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(17700,T)  
FT /\*tag= m  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(18074, TC)  
FT /\*tag= n  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(19328,T)  
FT /\*tag= o  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(19570,G)  
FT /\*tag= p  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(20892,T)  
FT /\*tag= q  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT 21555..21615  
FT exon 21555..21615  
FT /\*tag= r  
FT intron 21616..22462  
FT /\*tag= s  
FT exon 22463..22523  
FT /\*tag= t  
FT intron 22524..22974  
FT /\*tag= u  
FT exon 22975..23052  
FT /\*tag= v  
FT intron 23053..23711  
FT /\*tag= w  
FT exon 23712..23801  
FT /\*tag= x  
FT intron 23802..25392  
FT /\*tag= y  
FT exon 25393..25458  
FT /\*tag= z  
FT intron 25459..25613  
FT /\*tag= aa  
FT exon 25614..25769  
FT /\*tag= ab





Qy	1613	ACCAAGCCCTGTGGGTACTCCAGGGGCAAGAGAGCCCTGGGGTCAGTGAACATGTGTCAAG	1672
Db	26702	ACCAAGCCCTGTGGGTACTCCAGGGGCAAGAGAGCCCTGGGGTCAGTGAACATGTGTCAAG	26761
Qy	1673	CCAAACCATGACATCCACAAAGGGGAGCATTTGGAAATGAAGCATAGACTCTCTATGTATCA	1732
Db	26762	CCAAACCATGACATCCACAAAGGGGAGCATTTGGAAATGAAGCATAGACTCTCTATGTATCA	26821
Qy	1733	GGTTAAGAGCAAGGGAGAGCTGCCAGGACACAGCATTTTGACACGACAGAGGGGAATGTAG	1792
Db	26822	GGTTAAGAGCAAGGGAGAGCTGCCAGGACACAGCATTTTGACACGACAGAGGGGAATGTAG	26881
Qy	1793	CAACAGCAGGGCCCTCTAGGCCCCATCTTTCCATTTCTTAGGTAAAGAGCATTTTCTCTCA	1852
Db	26882	CAACAGCAGGGCCCTCTAGGCCCCATCTTTCCATTTCTTAGGTAAAGAGCATTTTCTCTCA	26941
Qy	1853	GACTCCAGCGGAGAGCATGAGGCTTAGGCTTCAGCAACCAAGGTTCTCTGGAGCCCAAA	1912
Db	26942	GACTCCAGCGGAGAGCATGAGGCTTAGGCTTCAGCAACCAAGGTTCTCTGGAGCCCAAA	27001
Qy	1913	GTTTATGGGAGAGGCGCAAGACTTCATCGGAGAGAGCAAGCGCCCTGGGTAGAAC	1972
Db	27002	GTTTATGGGAGAGGCGCAAGACTTCATCGGAGAGAGCAAGCGCCCTGGGTAGAAC	27061
Qy	1973	GCTTGTGTCTGTCTCTTTGGCCCTTTAAGACAAAGCGCTCATCTTTGCCCTCTTACCTCCTG	2032
Db	27062	GCTTGTGTCTGTCTCTTTGGCCCTTTAAGACAAAGCGCTCATCTTTGCCCTCTTACCTCCTG	27121
Qy	2033	ATAGGCTTAGGCTTTGCCCAACACACTGTGGCTACAGGTGAGGGAAGAGACTCTCTTC	2092
Db	27122	ATAGGCTTAGGCTTTGCCCAACACACTGTGGCTACAGGTGAGGGAAGAGACTCTCTTC	27181
Qy	2093	CTCCAGAGTCTATGTTCCAGGAAGTTTCTTTAAACCCCATATGCCCAAGAGTAGTCTGTA	2152
Db	27182	CTCCAGAGTCTATGTTCCAGGAAGTTTCTTTAAACCCCATATGCCCAAGAGTAGTCTGTA	27241
Qy	2153	GGAGGCCCTTTAAGACGGAAACAAAGTAAATTACAGTCTCTACTGGGTTCTTGCCACCG	2212
Db	27242	GGAGGCCCTTTAAGACGGAAACAAAGTAAATTACAGTCTCTACTGGGTTCTTGCCACCG	27301
Qy	2213	TCCCAAGTGGGCGAGCCCTAGGAAGGGTCAATCTTAAAGCCACACATTAGCTGCACTG	2272
Db	27302	TCCCAAGTGGGCGAGCCCTAGGAAGGGTCAATCTTAAAGCCACACATTAGCTGCACTG	27361
Qy	2273	CGTGGCTGACGCCAAACAAAGAACTGGGTGTGTAGTATTGATCACTAAGAACCAAAAT	2332
Db	27362	CGTGGCTGACGCCAAACAAAGAACTGGGTGTGTAGTATTGATCACTAAGAACCAAAAT	27421
Qy	2333	CCAGGGCAGCTCATATGTGAAGGATAAGAACCTCACCTTCTTACTCTCCAAAGAAGTGT	2392
Db	27422	CCAGGGCAGCTCATATGTGAAGGATAAGAACCTCACCTTCTTACTCTCCAAAGAAGTGT	27481
Qy	2393	GGGAAAGAACCATCAAACTTCTCTCTGATTTACCAAAACAGGAAACACAGAGAGG	2452
Db	27482	GGGAAAGAACCATCAAACTTCTCTCTGATTTACCAAAACAGGAAACACAGAGAGG	27541
Qy	2453	GTGGCTCAGGACTTAGGGAACAGGGTATAGCTTAGGTGGAAGCAAGAGGAGCAGGA	2512
Db	27542	GTGGCTCAGGACTTAGGGAACAGGGTATAGCTTAGGTGGAAGCAAGAGGAGCAGGA	27601
Qy	2513	AGTTGTAAATCAGCTGGCTAATGAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGAC	2572
Db	27602	AGTTGTAAATCAGCTGGCTAATGAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGAC	27661
Qy	2573	TAGGCTGGAGTTGCTTCTTTGAAGATGGGACTCTTTGGGTATCAAGACCTATGCCACATC	2632
Db	27662	TAGGCTGGAGTTGCTTCTTTGAAGATGGGACTCTTTGGGTATCAAGACCTATGCCACATC	27721
Qy	2633	ACACTGGGGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTGTGTCTTACGCCAGGACTT	2692
Db	27722	ACACTGGGGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTGTGTCTTACGCCAGGACTT	27781

[illegible]

PR 30-MAR-2000; 2000US-00540763.  
XX (CURA-) CURAGEN CORP.  
XX Shimkets RA, Leach M;  
XX WPI; 2000-602362/57.  
DR P-PSDB; AAB41604.  
XX  
PT Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease.  
XX  
PS Claim 5; Page 1978-1979; 5507pp; English.  
XX  
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORF open reading frames 1 to 3161. The ORF  
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
CC antiproliferative; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORF-associated disorder. The nucleic acids can be used to express ORF  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease, to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
SQ Sequence 1666 BP; 333 A; 529 C; 461 G; 341 T; 0 U; 2 Other;  
  
Query Match 48.1%; Score 1568.2; DB 3; Length 1666;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1598; Conservative 0; Mismatches 3; Indels 18; Gaps 1;  
  
Qy 2 GCCCGCTGCCGCCCGCAGTTCCTCCGGCCCGCCCTGCGCCAGTCATGCGGAGCAGTACGA 61  
Db 48 GCCCGCTGCCGCCCGCAGTTCCTCCGGCCCGCCCTGCGCCAGTCATGCGGAGCAGTACGA 107  
  
Qy 62 TGTGCTGTTCGGCTGTCTGTGATCGGGGACTCCGGGGGTGGGGAAGACCTGCGCTGTGTG 121  
Db 108 TGTGCTGTTCGGCTGTCTGTGATCGGGGACTCCGGGGGTGGGGAAGACCTGCGCTGTGTG 167  
  
Qy 122 CCGCTTCCACGACACGAGTTCCTCCTCCGACATCTCCACCATCGGTGTTGACTTTAA 181  
Db 168 CCGCTTCCACGACACGAGTTCCTCCTCCGACATCTCCACCATCGGTGTTGACTTTAA 227  
  
Qy 182 GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACATCGCAGG 241  
Db 228 GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACATCGCAGG 287  
  
Qy 242 GCAGGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCAGGGGATATTTTT 301  
Db 288 GCAGGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCAGGGGATATTTTT 347  
  
Qy 302 GGTCTATGACATTAGCAGCGCGCTTACCAGCATCATGATGAGTGGGTCACTGACGT 361  
Db 348 GGTCTATGACATTAGCAGCGCGCTTACCAGCATCATGATGAGTGGGTCACTGACGT 407  
  
Qy 362 GGATGAGTACGACACGAGCGCTCCAGAGATCCTTATGGGAATAGGCTGATGAGGA 421  
Db 408 GGATGAGTACGACACGAGCGCTCCAGAGATCCTTATGGGAATAGGCTGATGAGGA 467  
  
Qy 422 GCAGAAACGGCAGGTGGGGAAGAGAGCAGGGCAG-----CAGCTGGC 463  
Db 468 GCAGAAACGGCAGGTGGGGAAGAGAGCAGGGCAG-----CAGCTGGC 527

Qy 464 GAAGGAGTATGGCATGACATTTCTATGAAACAAAGTGCCTGCACCAACCTCAACATTAAAGA 523  
Db 528 GAAGGAGTATGGCATGACATTTCTATGAAACAAAGTGCCTGCACCAACCTCAACATTAAAGA 587  
  
Qy 524 GTCAATTCACGGCTGTGACAGAGCTGGTGTGACGGCCCATAGGAAGAGAGCTGGAGGGCT 583  
Db 588 GTCAATTCACGGCTGTGACAGAGCTGGTGTGACGGCCCATAGGAAGAGAGCTGGAGGGCT 647  
  
Qy 584 CCGGATCGGTGCCAGCAATGAGTTGGCACTGGGAGAGCTGGAGAGGAGGGGCAACACC 643  
Db 648 CCGGATCGGTGCCAGCAATGAGTTGGCACTGGGAGAGCTGGAGAGGAGGGGCAACACC 707  
  
Qy 644 CGAGGCGCCGACGCAACTCTTCGAAAAACCTGCTGTGTGTGAGTCTCTGTGTGGGGCACCCCA 703  
Db 708 CGAGGCGCCGACGCAACTCTTCGAAAAACCTGCTGTGTGTGAGTCTCTGTGTGGGGCACCCCA 767  
  
Qy 704 CACGACACCCCTCTTCCCTCAGGAGGCGCTGGGCGACAGAGGGAGAGCGGGGGCTTTGGCC 763  
Db 768 CACGACACCCCTCTTCCCTCAGGAGGCGCTGGGCGACAGAGGGAGAGCGGGGGCTTTGGCC 827  
  
Qy 764 TGTGCTGTCTCTCGGTGATGACCTATTGAGTATCAGTAGCCTACTTCCCTCCCTGCC 823  
Db 828 TGTGCTGTCTCTCGGTGATGACCTATTGAGTATCAGTAGCCTACTTCCCTCCCTGCC 887  
  
Qy 824 TGGCCCTGAGAGCGGCTCTGCTGTCATCTCAAGCAGCCCTGTCCCGACCCCGTCCACCC 883  
Db 888 TGGCCCTGAGAGCGGCTCTGCTGTCATCTCAAGCAGCCCTGTCCCGACCCCGTCCACCC 947  
  
Qy 884 TGGAGTGGTCTTTCTTACGCTGTTTCCCGACCCAGGCGCTGTACGACCCCGCAGTGT 943  
Db 948 TGGAGTGGTCTTTCTTACGCTGTTTCCCGACCCAGGCGCTGTACGACCCCGCAGTGT 1007  
  
Qy 944 GCAGCAGCAGTGTCTCACCATCCGACCCAGCAGCAGCAGCAGGCGCTGGAGTCAG 1003  
Db 1008 GCAGCAGCAGTGTCTCACCATCCGACCCAGCAGCAGCAGCAGGCGCTGGAGTCAG 1067  
  
Qy 1004 GCACATTTTACGCTGCTCTTCTCCGTGATGCTGCTCTCTCTGCTTTTCTCTCTCTTC 1063  
Db 1068 GCACATTTTACGCTGCTCTTCTCCGTGATGCTGCTCTCTCTGCTTTTCTCTCTCTTC 1127  
  
Qy 1064 CCCCCTTCTCTTCTCTGACCCCTCCCTCCCGTGGTTCGATCAAGCTCCTCAAA 1123  
Db 1128 CCCCCTTCTCTTCTCTGACCCCTCCCTCCCGTGGTTCGATCAAGCTCCTCAAA 1187  
  
Qy 1124 CCCCCTTCCCGGTGCTGCTGTGTGACGCTGCTCTTCTCTCTCTCTCTCTCTCTCTCT 1183  
Db 1188 CCCCCTTCCCGGTGCTGCTGTGTGACGCTGCTCTTCTCTCTCTCTCTCTCTCTCTCT 1247  
  
Qy 1184 CAAGGGGATGGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTTCC 1243  
Db 1248 CAAGGGGATGGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTTCC 1307  
  
Qy 1244 ACCCTGCTCTGTGGGTGGGCCAAAGGCTACAGGCTGCTTCTCTCTTCCCGACCCCA 1303  
Db 1308 ACCCTGCTCTGTGGGTGGGCCAAAGGCTACAGGCTGCTTCTCTCTTCCCGACCCCA 1367  
  
Qy 1304 CTGTCCCTCATGTGCGCATGGGCTGCTCCCTCCAGTACCTCGGAAAGTGGAGCATCGAGG 1363  
Db 1368 CTGTCCCTCATGTGCGCATGGGCTGCTCCCTCCAGTACCTCGGAAAGTGGAGCATCGAGG 1427  
  
Qy 1364 TAGGAGGGAACAGCAACCGGGGAGTCTCGAGCTTGGGGCTGCCCTTACCTTACCCATT 1423  
Db 1428 TAGGAGGGAACAGCAACCGGGGAGTCTCGAGCTTGGGGCTGCCCTTACCTTACCCATT 1487  
  
Qy 1424 CCCCAGCAGAGCTTTTGGCCCTTGTCTGGCTGCCCGCTGCTCTTCTTTGGGGAACTGAGCTC 1483  
Db 1488 CCCCAGCAGAGCTTTTGGCCCTTGTCTGGCTGCCCGCTGCTCTTCTTTGGGGAACTGAGCTC 1547  
  
Qy 1484 AGAGGCGAGTGTCTTTCAGAGAGGAAAAAATAATAGGGGTGCGAGGATATAAAGTCACT 1543  
Db 1548 AGAGGCGAGTGTCTTTCAGAGAGGAAAAAATAATAGGGGTGCGAGGATATAAAGTCACT 1607











QY 3078 GCAGCCACCTGCTCCTGAGCCTCCATCTATCTCCCTGTGACGGGTGAACCTTGGTACTG 3137  
 DB 6122 ACAGCCACCTACTCTCTAAACCTCCACTATCTCCCTATAAGTAACATTCCTGATCTA 6063  
 QY 3138 TGCTCGGGTCCATATATGAATGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTA 3197  
 DB 6062 TAICTCGAATCCATATATTAATTAACAAATTCATCTATTTTAAACACAAATATTTA 6003  
 QY 3198 CAAATTAAGATTTTCAACACCCAAACAAAACAAAACAAAACAAAACAAAACAAA 3257  
 DB 6002 CAAATTAAGATTTTCAACACCCAAACAAAACAAAACAAAACAAAACAAAACAAA 5943

RESULT 8  
 ID ABL34139/c  
 AC ABL34139;  
 XX 26-MAR-2002 (first entry)  
 DT Human immune system associated gene SEQ ID NO: 2112.  
 DE Human; immune system disease; cytosine methylation; antiasthmatic;  
 XX antiarteriosclerotic; antianemic; cytostatic; nootropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antithyroid; antidiabetic; antipeoriatic;  
 KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
 KW ds.  
 XX Homo sapiens.  
 OS WO200200928-A2.  
 XX 03-JAN-2002.  
 PD 02-JUL-2001; 2001WO-EP007537.  
 XX 30-JUN-2000; 2000DE-01032529.  
 PR 01-SEP-2000; 2000DE-01043826.  
 XX (EPIG-) EPIGENOMICS AG.  
 PA Olek A, Piepenbrock C, Berlin K;  
 XX WPI; 2002-130909/17.  
 XX Nucleic acid comprising fragment of chemically modified gene, useful for  
 PT diagnosis and treatment of diseases associated with abnormal cytosine  
 PT methylation.  
 PS Claim 1; SEQ ID NO 2112; 32pp + Sequence Listing; German.  
 XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention  
 XX Sequence 7924 BP; 1973 A; 83 C; 1827 G; 4041 T; 0 U; 0 Other;  
 SQ Query Match 35.6%; Score 1160.8; DB 6; Length 7924;  
 Best Local Similarity 74.1%; Pred No. 1.3e-271;  
 Matches 1468; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

QY 1278 TGCTTCTCTCTCCCAACCCCACTGCTCCCTCATGTGCAATGGCGTGGCTCCCGAG 1337  
 DB 7922 TACTTCTCTCTCTCCCAACCCCACTGCTCCCTCATATATACCAATTAACCTACCTCCCA 7863

QY 1338 TGACCTGCGAAAGTGGAGCATCGAGTAGGAGGAAACAGCAACCGGGAGTCTCTGAGC 1397  
 DB 7862 TAACTACGAAATTAACATCGAATTAACAAACGCAACCAAAATCTCTGAAAC 7803  
 QY 1398 CTGGGCTGCTCCTTACCTTACCCATTTCCCGACGAGAGCTTTGGCTTGGCTGCCCC 1457  
 DB 7802 CTAAAACTACCTTACCTTACCGCATTTCCCGACCAAACTTTTACCTTTAACTTACCC 7743  
 QY 1458 GCTGCTCTTTGGGAACTGAGCTCAGAGGAGGTGCTTTCAGAGAGGAAACAAATGA 1517  
 DB 7742 GCTTACCTTTTAAAAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 7683  
 QY 1518 GGGTGGCAGGATTAAGATGACCTTCACTTCTTACCTTCCATTCCTTACCTTCCATTC 1577  
 DB 7682 AATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7623  
 QY 1578 TTCTCTCCACCTGGCTCCCAAAATTTAAAGATGTGGACCAAGGCTGTGGGTACTCCAGG 1637  
 DB 7622 TTCTCTCCACCTAACTTCCCAAAATTTAAATAATAAATAAATAAATAAATAAATAAATA 7563  
 QY 1638 GAGGAGAGCCCTGGGGTCACTGACACTGTGAGGCAACCACTGCACTCCACAAAGGGA 1697  
 DB 7562 AAAAAAAACCTTAAATCAATTAACACTTAACTTAACTTAACTTAACTTAACTTAACT 7503  
 QY 1698 GCATTTGGAAATGAAGGACTAGCTTCTTATGTATGAGTTAAGAGCAAGGAGAGCTGGCC 1757  
 DB 7502 ACATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7443  
 QY 1758 AGGACAGAGTGTTCACAGCAGAGGAGGAGTGTAGCAACAGAGGAGGCTTCTAGGCCCA 1817  
 DB 7442 AAAAAACAATTTTACACAAACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 7383  
 QY 1818 TCTTCTATTTCTAGTGAAGAGACTTTTCTCAGACTCCAGGCGGAGGACTGAGCCT 1877  
 DB 7382 TCTTCTATTTCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7323  
 QY 1878 AGCTTTCAGCAACCAAGGTTCTCTGGGACCCAAAGTTTATGGGAGAGGAGGCAAGACTT 1937  
 DB 7322 AACTTTCACCAACCAAAATTTCTTAAACCCCAAAATTTTAAATAAATAAATAAATAAATA 7263  
 QY 1938 CATGGGAGAGAGAGAGGAGGAGGCTGCTGAGTGAACGCTTGGTCTGCTTCTTGGCTT 1997  
 DB 7262 CATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 7203  
 QY 1998 TAAGACAAAGCGCTCATCTTGGCTTCTTACCTCTGATAGGCTTGGAGGTTTGGCAACAC 2057  
 DB 7202 TAAACCAAAAGCTCATCTTACCTCTTACCTCTTACCTCTTACCTCTTACCTCTTACCTCT 7143  
 QY 2058 ACTGTGCTTACAGTGGAGGAGAGGAGTCTTCTTCTTCCAGAGTGTATGTTTACGGAAGT 2117  
 DB 7142 ACTATTAACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 7083  
 QY 2118 TCTTTTAAACCCATATGAGGCTTCTGAGGAGGAGGCTTCTTAAAGACGGAACCAAG 2177  
 DB 7082 TCTTTTAAACCCAT 7023  
 QY 2178 TAAATTTACAGTCTTCTGAGGCTTCTTCCGACCGTCCCAAGGTTGGGAGGCTTAGGAA 2237  
 DB 7022 TAAATTTACCAATTTCTTAAATTTCTTACCCACCGTCCCAAAATAAAGCAACCTTAAATA 6963  
 QY 2238 GAGGCTCATTTCTTAAAGCCACATTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2297  
 DB 6962 AAAAAATCATTTTAAACCAACATTAACACTTACCTTACCTTACCTTACCTTACCTTACCT 6903  
 QY 2298 TGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGCACTCATATGTGAGGATA 2357  
 DB 6902 TAAATATTAATAATTTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 6843  
 QY 2358 AGACCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2417  
 DB 6842 AAAAACTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6783









CC amount of expression of the polypeptide in a biological sample and  
CC diagnosing a pathological condition or a susceptibility to a pathological  
CC condition based on the presence or amount of expression of the  
CC polypeptide. The polypeptide is also useful for identifying a binding  
CC partner to the polypeptide, which involves contacting the polypeptide  
CC with a binding partner and determining whether the binding partner  
CC effects an activity of the polypeptide. The polypeptide or the nucleic  
CC acid encoding the polypeptide is useful for preventing, treating, or  
CC ameliorating a medical condition, which involves administering the  
CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid  
CC is useful for diagnosing a pathological condition or a susceptibility to  
CC a pathological condition in a subject, which involves determining the  
CC presence or absence of a mutation in the nucleic acid, and diagnosing a  
CC pathological condition or susceptibility to a pathological condition  
CC based on the presence or absence of the mutation. The polypeptide, the  
CC nucleic acid and an antibody to the polypeptide are useful for treating  
CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal  
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,  
CC the nucleic acid and the antibody are useful as immunosuppressive agents,  
CC as adjuvants to enhance immune responses, and as agents to induce higher  
CC affinity antibodies and increase serum immunoglobulin concentrations. The  
CC present sequence represents cDNA encoding a novel human protein. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format direct from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20020168711.  
XX  
SQ Sequence 2021 BP; 464 A; 564 C; 553 G; 440 T; 0 U; 0 Other;

Query Match 35.6%; Score 1160.6; DB 9; Length 2021;  
Best Local Similarity 99.2%; Pred. No. 7.9e-272;  
Matches 1166; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 2 GCCCGCTCCCGCCGCGAGTTCGCCGCGCCGCTGGCCCCAGTCATCGGAGCAGTACGA 61  
DB 38 GCCCGCTCCCGCCGCGAGTTCGCCGCGCCGCTGGCCCCAGTCATCGGAGCAGTACGA 97  
QY 62 TGTGCTGTTCGGCTGTGCTGTGATCGGGGACTCCGGGGTGGGCAAGACCTGCCCTGCTGTG 121  
DB 98 TGTGCTGTTCGGCTGTGCTGTGATCGGGGACTCCGGGGTGGGCAAGACCTGCCCTGCTGTG 157  
QY 122 CCGCTTCACCGACAAACAGTTCCTCTCGCACATCTCCACCATCGGTGTGATTTAA 181  
DB 158 CCGCTTCACCGACAAACAGTTCCTCTCGCACATCTCCACCATCGGTGTGATTTAA 217  
QY 182 GATGAACACCATAGAGTACGAGCATCAAGTCGGATACAGATCTGGGACACTGCAGG 241  
DB 218 GATGAACACCATAGAGTACGAGCATCAAGTCGGATACAGATCTGGGACACTGCAGG 277  
QY 242 GCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGGCGGCGCCAGGGGATATTTT 301  
DB 278 GCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGGCGGCGCCAGGGGATATTTT 337  
QY 302 GGTCTATGACATTAGCAGCAGCGCTCTTACAGCAATCATGAAAGTGGGTGAGTACGT 361  
DB 338 GGTCTATGACATTAGCAGCAGCGCTCTTACAGCAATCATGAAAGTGGGTGAGTACGT 397  
QY 362 GGATGACTACGACAGAGCGCTCCAGAGATCTTTATGGGATTAAGCTGATGAGGA 421  
DB 398 GGATGACTACGACAGAGCGCTCCAGAGATCTTTATGGGATTAAGCTGATGAGGA 457  
QY 422 GCAGAAACGGCAGGTGGGAGAGAGAGCAGGACAGCTGGCGAAGGAGTATGGCATGGA 481  
DB 458 GCAGAAACGGCAGGTGGGAGAGAGAGCAGGACAGCTGGCGAAGGAGTATGGCATGGA 517  
QY 482 CTTCTATGAAACAAGTGCCTGACCAACCTCAACATTAAAGAGTCAATTCACGGCTGTGAC 541  
DB 518 CTTCTATGAAACAAGTGCCTGACCAACCTCAACATTAAAGAGTCAATTCACGGCTGTGAC 577  
QY 542 AGAGCTGTGCTGACGCCCATAGAGAGAGCTGGAGAGGCTCCGGATGGTGGCCAGCA 601  
DB 578 AGAGCTGTGCTGACGCCCATAGAGAGAGCTGGAGAGGCTCCGGATGGTGGCCAGCA 637  
QY 602 TGAGTTGGCACTGGCAGAGCTGGCAGGAGGAGGAGGCAAAACCCGAGGCGCCGAGCACTC 661

DB 638 TGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGCAAAACCCGAGGCGCCGAGCACTC 697  
QY 662 TTGCAAAAACCTGCTGTGTGCTGAGTCTGTGTGTGGGAGCAACCCACACAGACACCCCTCTTCCC 721  
DB 698 TTGCAAAAACCTGCTGTGTGCTGCGTCTGTGTGGGGGGACACCCACACAGACACCCCTCTTCCC 757  
QY 722 TCAGGAGGCGGTGGGCGACAGCGGAGCGGGGCTTTGGCCCTGCTGTCTCTCTCTGTG 781  
DB 758 TCAGGAGGCGGTGGGCGACAGCGGAGCGGGGCTTTGGCCCTGCTGTCTCTCTCTGTG 817  
QY 782 TGATGACCCCTATTGAGTATCAGTAGCCACTACTCCCTCTGCTGGCCCTGAGAGCGGCTC 841  
DB 818 TGATGACCCCTATTGAGTATCAGAAGCCACTACTCCCTCTGCTGGCCCTGAGAGCGGCTC 877  
QY 842 TGCTGTCTATCTAAGCAGCCCTGTGTCAGCCGCTGTCACCCCTGGAGTGTGTTCTTCTCAG 901  
DB 878 TGCTGTCTATCTAAGCAGCCCTGTGTCAGCCGCTGTCACCCCTGGAGTGGGCTTCTTCTCAG 937  
QY 902 CCGTGTTCCTCCAGCCAGGCTGTGACGACCCCGGATGTCGCGCAAGCACTGTCTCA 961  
DB 938 CCGTGTTCCTCCAGCCAGGCTGTGACGACCCCGGATGTCGCGCAAGCACTGTCTCA 997  
QY 962 CCATCCCGCACCCACAGACAGCCAGGCTGGAGTCCAGGCCACTTTTCAGCTGTCTCC 1021  
DB 998 CCATCCCGCACCCACAGACAGCCAGGCTGGAGTCCAGGCCACTTTTCAGCTGTCTCC 1057  
QY 1022 TTTCTCGTGTCATGCTGTCT 1081  
DB 1058 TTTCTCGTGTCATGCTGTCT 1117  
QY 1082 GACCCCTCCCTCCGGTGGTTCGTATCAAGCTCTCAAAACCCCGTCCCGGTCGTC 1141  
DB 1118 GACCCCTCCCTCCGGTGGTTCGTATCAAGCTCTCTCAAAACCCCGTCCCGGTCGTC 1177  
QY 1142 CTGCTGTGTGTCAGCTGCT 1176  
DB 1178 CTGCTGTGTGTCAGCTGCT 1212  
RESULT 11  
ABQ93361  
ID ABQ93361 standard; cDNA; 1054 BP.  
XX AC ABQ93361;  
XX AC ABQ93361;  
DT 14-OCT-2002 (first entry)  
XX Human cDNA SEQ ID NO 74.  
XX Human; vulnary; dermatological; neuroprotective; nootropic; cancer;  
XX antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;  
XX antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;  
XX burn; central nervous system disorder; Alzheimer's disease;  
XX Parkinson's disease; Huntington's disease; immune disorder;  
XX autoimmune disorder; multiple sclerosis; diabetes; allergy; gene; ss.  
OS Homo sapiens.  
XX WO200218424-A2.  
XX 07-MAR-2002.  
XX 31-AUG-2001; 2001WO-US027093.  
XX 01-SEP-2000; 2000US-00654935.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;  
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;  
XX WPI; 2002-583321/62.  
DR





```

PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
DI
XX WPI; 2001-465460/50.
XX P-PSDB; AAU17555.
DR
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders.
XX
XX Claim 1; SEQ ID NO 507; 880pp; English.
XX
XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative disorders
CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities
CC (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),
CC respiratory disorders, dermatological disorders, in wound healing,
CC epithelial cell proliferation, endocrine disorders (e.g. Addison's
CC disease), reproductive system disorders, gastrointestinal disorder
CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of
CC B-cell responsiveness to pathogens, activators of T-cells, to induce
CC higher affinity antibodies, and as a means to induce tumour proliferation
CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-
CC AAS27850 represent novel signal transduction pathway protein coding
CC sequences and PCR primers of the invention
XX

Query Match 16.4%; Score 534.2; DB 4; Length 566;
Best Local Similarity 98.9%; Pred. No. 1.2e-119;
Matches 533; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GCCCGCTGCCCGCCGCGAGTTCCCGGCCCGCCGCTGCGCCAGTCATGGCAAGCAGTAGCA 61
Db 28 GCCCGCTGCCCGCCGCGAGTTCCCGGCCCGCCGCTGCGCCAGTCATGGCAAGCAGTAGCA 87

Qy 62 TGTGCTGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTCCCTGCTGTG 121
Db 88 TGTGCTGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTCCCTGCTGTG 147

Qy 122 CCGCTTCCCGACACGAGTTCACCTCTCGCACATCTCCACCACCGTGTGACTTTAA 181
Db 148 CCGCTTCCCGACACGAGTTCACCTCTCGCACATCTCCACCACCGTGTGACTTTAA 207
```



Qy 182 GATGAGACCATAGAGTACAGCGCATCAAAAGTCCGATACAGATCTGGGACACTGCAGG 241  
Db |||||||  
Qy 208 GATGAGACCATAGAGTACAGCGCATCAAAAGTCCGATACAGATCTGGGACACTGCAGG 267  
Db |||||||  
Qy 242 GCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGCGGGCCCCAGGGGATATTTT 301  
Db |||||||  
Qy 268 GCAGGAGAGATACAGACCATCAAAAGCAGTACTATCGCGGGCCCCAGGGGATATTTT 327  
Db |||||||  
Qy 302 GGTCTATGATATAGCAGCAGCGCTTTACACGACATCATGAAGTGGTCACTGACGT 361  
Db |||||||  
Qy 328 GGTCTATGATATAGCAGCAGCGCTTTACACGACATCATGAAGTGGTCACTGACGT 387  
Db |||||||  
Qy 362 GGATGAGTACGACCAAGCGCTTCAGAGATCTTTATTTGGGAATTAAGCTCATGAGGA 421  
Db |||||||  
Qy 388 GGATGAGTACGACCAAGCGCTTCAGAGATCTTTATTTGGGAATTAAGCTCATGAGGA 447  
Db |||||||  
Qy 422 GCAGAAACGGCAGGTGGGAAGAGAGAGAGCGCAGCAGCTGGCGAAGAGTATGGCATGGA 481  
Db |||||||  
Qy 448 GCAGAAACGGCAGGTGGGAAGAGAGAGAGCGCAGCAGCTGGCGAAGAGTATGGCATGGA 507  
Db |||||||  
Qy 482 CTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAGAGTCACTACGGGTCTGA 540  
Db |||||||  
Qy 508 CTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAGAGTCACTACGGGTCTGA 566  
Db |||||||

RESULT 14

ADB93650 ADB93650 standard; cDNA; 566 BP.

XX AC ADB93650;

XX DT 04-DEC-2003 (first entry)

XX DE Human cDNA encoding a novel protein #497.

XX KW ss; Gene; human; autoimmune disease; Parkinson's disease; silicosis;

XX KW gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;

XX KW immunosuppressive agent; adjuvant; enhance immune response;

XX KW higher affinity antibody induction;

XX KW increased serum immunoglobulin concentration.

XX OS Homo sapiens.

XX PN US2002168711-A1.

XX PD 14-NOV-2002.

XX PF 17-JAN-2001; 2001US-00764868.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 22-AUG-2000; 2000US-0226868P.

XX PR 30-AUG-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229343P.

XX PR 01-SEP-2000; 2000US-0229344P.

XX PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244817P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-719985/68.

P-PSDB; ADB94263.

New isolated polypeptide useful for diagnosing and treating

immunosuppressive conditions such as autoimmune disease and Parkinson's

disease.

Claim 3; SEQ ID NO 507; 345pp; English.

The invention relates to an isolated polypeptide. The polypeptide is

useful for diagnosing a pathological condition or a susceptibility to a

pathological condition in a subject, by determining the presence or

amount of expression of the polypeptide in a biological sample and

diagnosing a pathological condition or a susceptibility to a pathological

condition based on the presence or amount of expression of the

polypeptide. The polypeptide is also useful for identifying a binding

partner to the polypeptide, which involves contacting the polypeptide

with a binding partner and determining whether the binding partner

affects an activity of the polypeptide. The polypeptide or the nucleic

acid encoding the polypeptide is useful for preventing, treating, or

ameliorating a medical condition, which involves administering the

polypeptide or the nucleic acid to a mammalian subject. The nucleic acid

is useful for diagnosing a pathological condition or a susceptibility to

a pathological condition in a subject, which involves determining the

presence or absence of a mutation in the nucleic acid, and diagnosing a

pathological condition or susceptibility to a pathological condition

based on the presence or absence of the mutation. The polypeptide, the

nucleic acid and an antibody to the polypeptide are useful for treating

autoimmune disease, Parkinson's disease, silicosis, gastrointestinal

disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,

the nucleic acid and the antibody are useful as immunosuppressive agents,

as adjuvants to enhance immune responses, and as agents to induce higher

affinity antibodies and increase serum immunoglobulin concentrations. The

present sequence represents cDNA encoding a novel human protein. Note:

The sequence data for this patent did not form part of the printed

specification but was obtained in electronic format direct from USPTO at

seqdata.uspto.gov/sequence.html?DocID=20020168711.

Sequence 566 BP; 140 A; 145 C; 172 G; 103 T; 0 U; 6 Other;

Query Match 16.4%; Score 534.2; DB 9; Length 566;  
Best Local Similarity 99.9%; Pred. No. 1.2e-119;  
Matches 533; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCCCGTCGCCGCCGAGTCCCGCGCCCGCTGGCCCGCATGTCGGGAGCAGTACGA 61  
DB 28 GCCCGTCGCCGCCGAGTCCCGCGCCCGCTGGCCCGCATGTCGGGAGCAGTACGA 87  
QY 62 TGTCTGTTCGGCTGTCTGTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTGCTGTG 121  
DB 88 TGTCTGTTCGGCTGTCTGTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTGCTGTG 147  
QY 122 CCGTTTACCGAACAAGAGTTCATCTCTGCGACATCTCCACCATCGGTGTGACTTTAA 181  
DB 148 CCGTTTACCGAACAAGAGTTCATCTCTGCGACATCTCCACCATCGGTGTGACTTTAA 207  
QY 182 GATGAAGACCATAGAGGTAGAGCGGCATCAAGTGGGATACAGATCTGGGACACTGCAGG 241  
DB 208 GATGAAGACCATAGAGGTAGAGCGGCATCAAGTGGGATACAGATCTGGGACACTGCAGG 267  
QY 242 GCAGGAGAGATACAGACCATCAACAAGCAGTACTATCGGCGGCCCGAGGGATATTTT 301  
DB 268 GCAGGAGAGATACAGACCATCAACAAGCAGTACTATCGGCGGCCCGAGGGATATTTT 327  
QY 302 GGTCTATGACATTAGCAGCGAGCGCTCTTACCAGACATCATGAAGTGGTCACTGACGT 361  
DB 328 GGTCTATGACATTAGCAGCGAGCGCTCTTACCAGACATCATGAAGTGGTCACTGACGT 387  
QY 362 GGATGATACGACAGAGCGTCCAGAGATCTTATTTGGGATTAAGGCTGATGAGGA 421  
DB 388 GGATGATACGACAGAGCGTCCAGAGATCTTATTTGGGATTAAGGCTGATGAGGA 447  
QY 422 GCAGAAACGCGAGTGGGAGAGAGAGGAGGCGAGCAGCTGGCGAAGAGTATGGCATGGA 481  
DB 448 GCAGAAACGCGAGTGGGAGAGAGAGGAGGCGAGCAGCTGGCGAAGAGTATGGCATGGA 507  
QY 482 CTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAAGAGTCATTACGCGTCTGA 540  
DB 508 CTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAAGAGTCATTACGCGTCTGA 566

RESULT 15  
ABK44752  
ID ABK44752 standard; cDNA; 481 BP.  
XX AC ABK44752;  
XX AC  
XX AC  
XX 05-JUN-2002 (first entry)  
XX cDNA encoding colon tumour protein, SEQ ID No 303.  
XX Human; colon tumour; vaccine; colon cancer; immunogenic; immunotherapy;  
XX gene; ss.  
XX Homo sapiens.  
XX OS  
XX WO200212328-A2.  
XX PN  
XX PD  
XX 14-FEB-2002.  
XX 31-JUL-2001; 2001WO-US024218.  
XX 03-AUG-2000; 2000US-0223283P.  
XX 28-MAR-2001; 2001US-0279763P.  
XX 29-JUN-2001; 2001US-0302051P.  
XX (CORI-) CORIXA CORP.  
XX PI King GE, Meagher MJ, Xu J, Secretist H;  
XX WPI; 2002-241739/29.  
XX New colon cancer polypeptides and polynucleotides, useful as vaccines.

PT for diagnosing, preventing, and treating colon cancer, and as markers for  
PT the progression of cancer.  
XX Claim 1; SEQ ID NO 303; 147pp; English.  
XX The invention relates to polynucleotides encoding colon tumour proteins.  
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical  
CC compositions, such as vaccines, for the diagnosis, prevention, and  
CC treatment of colon cancer. Polynucleotide sequences may be used as  
CC hybridisation probes or primers, and in the design and preparation of  
CC ribozyme molecules for inhibiting expression of tumour polypeptides and  
CC proteins in tumour cells. The compositions are useful for stimulating an  
CC immune response against cancer, particularly for the immunotherapy of  
CC colon cancer, and as markers for the progression of cancer. ABK4450-  
CC ABK46237 represent coding sequences of human colon tumour proteins of the  
CC invention. Note: With the exception of SEQ ID No 1 and 2, the sequence  
CC data for this patent did not form part of the printed specification but  
CC was supplied by the European Patent Office  
XX SQ Sequence 481 BP; 111 A; 113 C; 130 G; 125 T; 0 U; 2 Other;  
Query Match 14.7%; Score 477.4; DB 6; Length 481;  
Best Local Similarity 99.4%; Pred. No. 7.7e-106;  
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2582 GTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCAGCTGGG 2641  
DB 1 GTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCAGCTGGG 60  
QY 2642 CTAGGGAAGTAGGTGATGCGAGCCCTCAAGTCTGTCTTTCAGCGAGGACTTGAGAAGTTA 2701  
DB 61 CTAGGGAAGTAGGTGATGCGAGCCCTCAAGTCTGTCTTTCAGCGAGGACTTGAGAAGTTA 120  
QY 2702 TATTGGGACGTGGCTCCAAATCTGTGGACCAAGTATTTTCAGCTTTTCCCTGAAGATCAGGCAG 2761  
DB 121 TATTGGGACGTGGCTCCAAATCTGTGGACCAAGTATTTTCAGCTTTTCCCTGAAGATCAGGCAG 180  
QY 2762 GGTGCCATTCAATGTCTTTCTCTAGCCCCCTCAGGAAAGAGGACTATATTGTACT 2821  
DB 181 GGTGCCATTCAATGTCTTTCTCTAGCCCCCTCAGGAAAGAGGACTATATTGTACT 240  
QY 2822 GTACCCCTAGGGGTTCTGGAAGGGGAAAACATGGAATCAGGATTCATAGACTGATAGGCC 2881  
DB 241 GTACCCCTAGGGGTTCTGGAAGGGGAAAACATGGAATCAGGATTCATAGACTGATAGGCC 300  
QY 2882 TATCCACAAGGGCCATGACCTGGGAAAGGATATGGGACGAGAGAAATGGGATTTTAG 2941  
DB 301 TATCCACAAGGGCCATGACCTGGGAAAGGATATGGGACGAGAGAAATGGGATTTTAG 360  
QY 2942 GGTGCAGCTACGCTCACCTTAACTTTTGGTGGCTGGGCGCATCTTTGAGGCCCCAGACT 3001  
DB 361 GGTGCAGCTACGCTCACCTTAACTTTTGGTGGCTGGGCGCATCTTTGAGGCCCCAGACT 420  
QY 3002 GTTAAGCAGGCTCTGCTGGCTGTTTACTCGTCAACCACTCTGCACTGCTGCTGTGAGA 3061  
DB 421 GTTAACCAAGCTCTGCTGGCTGTTTACTCGTCAACCACTCTGCACTGCTGCTGTGAGA 480  
QY 3062 C 3062  
DB 481 C 481

Search completed: March 21, 2004, 05:37:11  
Job time : 1158 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2004, 23:06:10 ; Search time 226 Seconds  
(without alignments)  
7997.677 Million cell updates/sec

Title: US-09-817-198C-1

Perfect score: 3257

Sequence: 1 tgcgcgtgccgcgcgcag.....aaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

.Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	174.2	5.3	730	4	US-09-300-958A-16
2	156.6	4.8	1340	2	US-08-824-873-2
3	156.6	4.8	1340	3	US-08-824-873-2
4	156	4.8	925	2	US-09-198-184-2
5	156	4.8	925	4	US-08-916-901-4
6	144.6	4.4	3936	4	US-09-154-602-4
7	138	4.2	842	4	US-09-919-172-49
8	136.2	4.2	875	4	US-09-255-920A-6
9	136.2	4.2	1106	4	US-09-075-454-10
10	136.2	4.2	2612	4	US-09-620-312D-959
11	133.4	4.1	1069	4	US-09-484-970B-142
12	124.2	3.8	723	4	US-09-620-312D-646
13	123.6	3.8	1102	4	US-09-016-434-1422
14	122.6	3.8	8137	4	US-09-566-921-7
15	120.6	3.7	847	2	US-08-773-423-4
16	115	3.5	833	4	US-09-620-312D-426
17	113	3.5	639	4	US-09-399-913-66
18	110	3.4	970	3	US-08-888-077A-28
19	106.2	3.3	1054	4	US-09-976-594-1096
20	102.2	3.1	3745	4	US-09-976-594-387
c 21	93.8	2.9	193	4	US-09-702-705-989
c 22	93.8	2.9	193	4	US-09-702-705-1161
c 23	93.8	2.9	193	4	US-09-736-457-989
c 24	93.8	2.9	193	4	US-09-736-457-1161
c 25	93.8	2.9	193	4	US-09-614-124B-989
c 26	93.8	2.9	193	4	US-09-614-124B-1161
c 27	93.8	2.9	193	4	US-09-671-325-989

28	93.8	2.9	193	4	US-09-671-325-1161	Sequence 1161, Ap
29	93.6	2.9	313	4	US-09-313-294A-7460	Sequence 7460, Ap
30	91.2	2.8	820	3	US-08-741-411-6	Sequence 6, Appli
31	91.2	2.8	820	4	US-09-016-434-112	Sequence 112, App
32	85.8	2.6	7218	1	US-08-232-463-14	Sequence 14, Appl
33	84	2.6	241	4	US-09-016-434-813	Sequence 813, App
34	81.2	2.5	921	4	US-09-016-434-1124	Sequence 1124, Ap
35	80.6	2.5	890	3	US-08-741-411-4	Sequence 4, Appli
36	80	2.5	1172	4	US-09-075-454-8	Sequence 8, Appli
37	80	2.5	1255	2	US-08-766-551-6	Sequence 6, Appli
38	80	2.5	1533	4	US-09-075-454-11	Sequence 11, Appl
39	80	2.5	1886	4	US-09-620-312D-647	Sequence 647, App
40	79.2	2.4	1984	4	US-09-023-655-7	Sequence 7, Appli
41	79	2.4	848	3	US-08-741-411-2	Sequence 2, Appli
42	78.8	2.4	912	4	US-09-016-434-1159	Sequence 1159, Ap
43	77.4	2.4	503	4	US-09-621-976-15288	Sequence 15288, A
44	77	2.4	1749	4	US-09-149-476-54	Sequence 54, Appl
45	75.6	2.3	1407	4	US-09-493-914-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-300-958A-16  
; Sequence 16, Application US/09300958A  
; Patent No. 6495319  
; GENERAL INFORMATION:  
; APPLICANT: McClelland, Michael  
; APPLICANT: Welsh, John  
; TITLE OF INVENTION: Reduced Complexity Nucleic Acid Targets and Methods of  
; TITLE OF INVENTION: Using Same  
; FILE REFERENCE: P-PH 3457  
; CURRENT APPLICATION NUMBER: US/09/300,958A  
; CURRENT FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/083,331  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: 60/098,070  
; PRIOR FILING DATE: 1998-08-27  
; PRIOR APPLICATION NUMBER: 60/118,624  
; PRIOR FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 730  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-300-958A-16

Query Match		5.3%;	Score 174.2;	DB 4;	Length 730;
Best Local Similarity		56.7%;	Pred. No. 4.9e-36;	Indels 3;	Gaps 1;
Matches 342;		Conservative 0;	Mismatches 258;		
Qy	7	CTGCCCCCGCCGAGTTC	CGCCGCGCTGCGCCGAGTCATCGCGAAGCAGTACGATGTC	66	
Db	44	CGGACCCCGCCGCGCGCGCGCTCGCCCGCCGCGCCGCGGACTACGACCACC	103		
Qy	67	TGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTGTCGCCGCT	126		
Db	104	TCTTCAAGCTGCTCATCATCGGCACAGCGGTGTGGCAAGCAGTCTTACTGTTCGTT	163		
Qy	127	TCACCGACAGGAGTTCACCTCTCCGACATCTCCACATCGGTGTTGACTTTAAGATGA	186		
Db	164	TTGACAGAACACTTCTTCAGGCGAGCTACATCACACGATCGGAGTGAATTCAGATCC	223		
Qy	187	AGACATAGAGGTAGAGCGGCATCAAGTGGCGGATACAGATCTGGGACACTGCAGGCGAGG	246		
Db	224	GGACCGTGGAGATCAACCGGGAGAGGTGAAGCTGCAGATCTGGGACACACGCGGGCAGG	283		
Qy	247	AGAGATACAGACCATCACAAAGCAGTACTATCGCGGGGCCACAGGGGATATTTTGTGCT	306		
Db	284	AGCGTTCGCGACCATCACTCCACGTATTATCGGGGAGCCACCGGGTCTATTGTGTTT	343		

QY 307 ATGACATTAGCAGGAGCGCTCTTACAGCACATCATGAAGTGGTTCAGTGCAGTGGATG 366  
Db 344 ACAGCGTCACACGTCGCGAGTCTCTTTGTCAACGTCAGCGGTGGCTTCACGAATCAACC 403  
QY 367 AGTACGACACAGAGGCGTCCAGAGATCTTATTGGGAATAAGGCTGATGAGGAGCAGA 426  
Db 404 AGAAC-----TGTGATGATGTGTGCGGAATATTAGTGGGTAAATAGATGACGACCCCTGAGC 460  
QY 427 AACCGCAGGTGGGAAGAGAGCAAGGCGAGCGAGTGGCGAAGAGATATGGCATGGACTTCT 486  
Db 461 GGAAGTGGTGGAGCGGAAGATGCCCTACAAATTCGCGGCGAGATGGGCATCCAGTTGT 520  
QY 487 ATGAAACAAGTGCCTGCACCAACCTCAACATTAAAGAGTCAATTCACCGCTGTACAGAGC 546  
Db 521 TCGAGACAGCGCCCAAGGAGAAATGTCAACGTGGAAGAGATGTTCACTGCATCACGAGC 580  
QY 547 TGGTCTGCAGGCCCATAGGAAGAGCTGGAAGCGCTCCGAGTGCCTGCGAGCAATAGT 606  
Db 581 TGGTCTCCGAGCAAGAAAGACAACTGGCAAAACAGCAGCAGCAACACAGAACGATG 640  
QY 607 TGG 609  
Db 641 TGG 643

RESULT 2  
US-08-824-873-2  
; Sequence 2, Application US/08824873  
; Patent No. 5843717  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,873  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1340 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PANCNOT04  
; CLONE: 738957  
US-08-824-873-2  
Query Match 4.8%; Score 156.6; DB 2; Length 1340;

Best Local Similarity 59.0%; Pred. No. 3e-31;  
Matches 305; Conservative 0; Mismatches 209; Indels 3; Gaps 2;  
QY 57 TAGGATGTCTGTTCCGGCTGTCTGATCGGGGACTCGGGGTGGGCAAGACCTGCCTG 116  
Db 22 TAGGAGCTCGCTTCAAGGTCTATGCTGGTGGGGGACTCGGGGTGGGGAAGACCTGTCTG 81  
QY 117 CTGTGC-CGCTTCAACGCAACGAGTTCCACTCTCTCGCACATCTCCACCATCGGTG--TT 173  
Db 82 CTGGGTGGCATTCAGGATGGTGTCTTCTGCGGGGACCTTCATCTCCACCGTAGCAT 141  
QY 174 GACTTTTAAGATGAAGACCATAGAGGTAGAGCGCATCAAAAGTCCGGATACAGATCTGGGAC 233  
Db 142 GACTTCCGGAACAAAGTTCTGGACGTGGATGTTGTGAAGGTGAAGCTGCAGATGTGGGAC 201  
QY 234 ACTGAGGCGAGAGAGATACAGACCATCAAAAGCAGTACTATCGGGGGGCCCCAGGG 293  
Db 202 ACAGCTGGTTCAGAGCGGTTCGCGCATGTGTACCCATGCTTACCGGATGCTCATGCT 261  
QY 294 ATATTTTTCGTCTATGACATTAGCAGCGCGCTCTTACGACACATCATGAAGTGGGTC 353  
Db 262 CTGCTGCTCTTACGATGTCAACAAAGGCTCTTTTGACAAACATCCAGGCTGGCTG 321  
QY 354 AGTGACGTGGATGAGTACGACCCAGAACGCGTCCAGAGATCTTATTGGGAATAAGGCT 413  
Db 322 ACCGAGATCCAGAGTACGCCAGCACGACGTGGCGCTCATGCTGTGGGGAACAAGGTG 381  
QY 414 GATGAGGAGCAAAACGCGAGGTGGGAAGAGCAGCAGCGCTGGCGAAGGAGTAT 473  
Db 382 GACTCTGCCCATGAGCGTGTGTGAAGAGGGAGGACGGGAGAGCTGGCCAAAGGAGTAT 441  
QY 474 GGCATGGACTTCTATGAACAAAGTGCCTGCACCAACCTCAACATTAAGAGTCAATTCAGC 533  
Db 442 GGAATGCCCTTCTATGAGACCGCGCCAGAGCGGCTCAACGTGGACTTGGCTTTCACA 501  
QY 534 CGTCTGACAGAGCTGTGTGTCAGGCGCCCATAGGAAGG 570  
Db 502 GCCATAGCAAGGAGTTGAAGCAGCGCTCCATGAAGG 538  
RESULT 3  
US-09-198-184-2  
; Sequence 2, Application US/09198184  
; Patent No. 6010859  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/198,184  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/824,873  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555		
; TELEFAX: 415-845-4166		
; INFORMATION FOR SEQ ID NO: 2:		
; SEQUENCE CHARACTERISTICS:		
; LENGTH: 1340 base pairs		
; TYPE: nucleic acid		
; STRANDEDNESS: single		
; TOPOLOGY: linear		
; IMMEDIATE SOURCE:		
; LIBRARY: PANCNOT04		
; CLONE: 738957		
US-09-198-184-2		
Query Match	4.8%; Score 156.6; DB 3; Length 1340;	
Best Local Similarity	59.0%; Pred. No. 3e-31;	
Matches 305; Conservative 0; Mismatches 209; Indels 3; Gaps 2		
QY	57 TAGATGTGCTTTCGGCTCCTCTCTGATCGGGAGACTCCGGGTGGGGAAGACCTGCCTG 116	
Db		
22 TACGACCTCGCTTTCAAGGTTCATGCTGTGTGGGGACTCGGGTGTGGGAAGACCTGTCTG 81		
QY	117 CTGTGC--CGTTTCACCGACAACGAGTTCCACTCTCTCGCACATCTCCACCATCGTG--TT 173	
Db		
82 CTGGTCCGATTCAAGGATGGTGTCTTCTCGCGGGGACCTTTCATCTCCACCGTAGCAT 141		
QY	174 GACTTTTAAGATGAAGACCATAGAGGTAGAAGGCATCAAAGTGGCGATACAGATCTGGGAC 233	
Db		
142 GACTTCGGAAACAAAGTTCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATGTGGAC 201		
QY	234 ACTCGAGGCGAGGAGAGATACGAGACCATCAAAAGCAGTAGTACTATCGCGGGGCCACAGGG 293	
Db		
202 ACAGCTGGTCAGAGACGGTTCCGAGTCTTACCATGCTACTACCGGGATGCTCATGCT 261		
QY	294 ATATTTTGTGCTATGACATTAGCAGCGAGCGCTCTTACGAGCACATCATGAAGTGGGTC 353	
Db		
262 CTGCTGCTGCTCTACGATGTCAACAAAGGCGCTCTTTTGACAAATCCAGGCTGGCGTG 321		
QY	354 AGTGACGTGGATGAGTACGACACCAAGGCGTCCAGAAGATCCCTTATTGGGAATAAGGCT 413	
Db		
322 ACCGAGATCCACGAGTAGCGCCACGACGACGTGGCGCTCATGCTGCTGGGGAACAAAGTG 381		
QY	414 GATGAGGAGCAGAAACGCGAGTCGGGAAGAGAGCAAGGGCAGCAGCTGCGCAAGAGGAT 473	
Db		
382 GACTCTGCCATGAGCGTGTGGTGAAGGGAGGACGGGGAGAGCTGCGCCAGAGGAT 441		
QY	474 GGCATGGACCTTCTATGAACAAAGTGCCTGCGACCAACCTTCAACATTTAAAGAGTCATTCAG 533	
Db		
442 GGACTGCCCTTCATGGAGACCAGCGCCCAAGACGGGCGCTCAACGTGGACTTGGCCTTCACA 501		
QY	534 CGTCTGACAGAGCTGGTGTGTCAGGCCCCAATAGGAAGG 570	
Db		
502 GCCATAGCAAAGGAGTTGAACAGACGGCTTCCATGAAGG 538		

```

RESULT 4
US-08-316-901-4
; Sequence 4, Application US/08916901
; Patent No. 5892012
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,901
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 925 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LIVRUT04
; CLONE: 2514506
;
US-08-916-901--4

Query Match 4.8%; Score 156; DB 2; Length 925;
Best Local Similarity 56.7%; Pred. No. 3.5e-31;
Matches 288; Conservative 0; Mismatches 220; Indels 0; Gaps 0

QY 36 GCCCCAGTCATCGGGAAGCAGTAGTCAGATGTCGTGTTCCGGCTGCTGCTGATCGGGACTCC 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 GCGCCGCCCATGAACCCCGAATATGACTACCTGTTTAAAGCTGCTTTTGATTGGCGACTCA 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 96 GGGGTGGGCAAGACCTCCCTGCTGTGCGCTTACCGCAACAGAGTTCCACTCTCTCGCAC 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 117 GCGTGGGCAAGTCATGCTGCTCTCTCGGTTGCTGATGACAGTACACAGAGAGTAC 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 156 ATCTCCACCATCGGTGTGATCTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAAGTG 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 ATCAGCACCATCGGGTGGACTTCAGATCCGAACCATCGAGCTGGATGGCAAAACTATC 236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 216 CGGATACAGATCTGGNACATCGCAGGCGCAGAGAGATACACAGACCATCACAAGCAGTAC 275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 AAATCTCAGATCTGGGACACAGCGGGCCAGAACGGTTCGGACCATCACTTCCAGTAC 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 276 TATCGCGGGGCCACAGGGGATATTTTTCGTCTATGACATTAGCAGCAGCGCTCTTTACCAG 335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 TACGGGGGCTCATGGCATCATCGTGGTGTATGACGTCACTGACAGGATCTTACGCC 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 336 CACATCATGAAGTGGGTCACTGACGTGGATGAGTACGCACAGAGGCGTCAGAAGATC 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 AACGTGAAGCAGTGGCTGTCAGGAGATTGACCGCTATGTCAGCGGAGAACGTCAATAAGCTC 416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 396 CTTATTGGGATTAAGCTGTATGAGGACAGAAACCGCAGGTGGAGAGAGACGACGGGCGAG 455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 417 CTGGTGGGCACAAAGACGACCTCACCCACAAAGAGGTGGTGGACCAACCAACGCCCAAG 476
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 456 CAGCTGCGCAAGGATATGGCATGGACTTCTATGAAACAAGTGCCTGCACCAACTCTCAAC 515
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 477 GAGTTTGCAGACTCTCTGGGCATCCCTCTTTGGAGACGACGCGCAAGATGCCACCAAT 536
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 516 ATTTAAGAGTCATTCACGCGTCTGACAG 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 GTCAGCAGCGGCTTCATGACCATGGCTG 564
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5  
US-09-154-602-4  
; Sequence 4, Application US/09154602  
; Patent No. 6300472





```

US-09-255-920A-6
; Sequence 6, Application US/09255920A
; Patent No. 6623980
; GENERAL INFORMATION:
; APPLICANT: Fisher, Joseph
; APPLICANT: Lorens, James
; APPLICANT: Anderson, David
; APPLICANT: Luo, Ying
; APPLICANT: Huang, Betty
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: EXO1 and EXO2, EXOCYTOTIC PROTEINS
; FILE REFERENCE: A65905-1/DJB/RMS
; CURRENT APPLICATION NUMBER: US/09/255,920A
; CURRENT FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: 60/075,534
; PRIOR FILING DATE: 1998-02-23
; PRIOR APPLICATION NUMBER: 60/086,650
; PRIOR FILING DATE: 1998-05-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3)
; OTHER INFORMATION: The n at position 3 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (6)..(8)
; OTHER INFORMATION: The n at positions 6 through 8 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (12)
; OTHER INFORMATION: The n at position 12 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (14)..(15)
; OTHER INFORMATION: The n at positions 14 through 15 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (17)..(18)
; OTHER INFORMATION: The n at positions 17 through 18 represents an
; OTHER INFORMATION: unknown.
; NAME/KEY: unsure
; LOCATION: (22)
; OTHER INFORMATION: The n at position 22 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (25)
; OTHER INFORMATION: The n at position 25 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (726)
; OTHER INFORMATION: The n at position 726 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (823)
; OTHER INFORMATION: The n at position 823 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (108)
; OTHER INFORMATION: The n at position 108 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (138)
; OTHER INFORMATION: The n at position 138 represents an unknown.
; NAME/KEY: unsure
; LOCATION: (204)
; OTHER INFORMATION: The n at position 204 represents an unknown.
; OTHER INFORMATION: The n at position 204 represents an unknown.
US-09-255-920A-6
Query Match 4.2%; Score 138; DB 4; Length 842;
Best Local Similarity 57.6%; Pred. No. 1.9e-26;
Matches 265; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

```

```

73 GGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTCTGCTGCGCTTCAACG 132
119 GGGTGATGCTTCTTGGAGANTCGGGCGTGGCAAAACCTGTTTCTGATCCAAATTTCAAG 178
133 ACAACGAGTTCCACTCCT---CGCACATCTCCACCATCGGTGTTGACTTTTAAGATGAGA 189
179 ACGGGGCTTCTGTCGCGAACCTTNTAGCCACCGTCGGCATAGACTTCAGGAATAAAG 238
190 CCATAGAGGTAGACGGCATCAAGTGGCGGATACAGATCTGGGACACTGCGAGGAGGAGA 249
239 TGGTGACAGTGGATGGTTCAGGGTGAAGCTTCAGATCTGGGACACTGCGAGGAGGAGC 298
250 GATACAGAGCATCAAAAGCAGTACTATCGCGGGCCCGGAGATATTTTTGCTTATG 309
299 GCTTCCGAGTGTGACCCATGCTTATTACCGAGATGCTCAGGCTTTGCTCTGTTGTATG 358
310 ACATTAGCAGGAGCGCTCTTACAGCAGATCATGAAGTGGGTGAGTGCAGTGCAGTGAAGT 369
359 ACATCAACCAACCACTCTCTTTTGACAAACATCAGGGCTGGCTCACAGAGATTATGAGT 418
370 AGCACCAGAGAGGCTCCAGAAGATCCTTATTGGGAATAAGGCTGATGAGGAGGAGAAAC 429
419 ATGCCAGAGGAGCTGGTATGATGCTTCTAGGCAACAGGCCGATGTAGAGCGGAA 478
430 GGCAGTGGGAAGAGAGCAAGGGCAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATG 489
479 GGGTGATCCGTTCTGAAGATGGAGAGACACTGGCCAGGGAATATGGTGTCTCTTTCATGG 538
490 AAACAAGTGGCTGCACCACCTCAACATTAAGAGATCATT 529
539 AGACAGTGGCAAGACTGGCATGAACGTGGAGTTGGCCTT 578

RESULT 8
US-09-075-454-10
; Sequence 10, Application US/09075454
; Patent No. 6391580
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; APPLICANT: Batra, Sajeev
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,454
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,551
; FILING DATE: DECEMBER 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555

```

TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 875 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: UCMCL5T01  
CLONE: 1528559  
US-09-075-454-10

Query Match 4.2%; Score 136.2; DB 4; Length 875;  
Best Local Similarity 57.5%; Pred. No. 5.7e-26;  
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

Qy 57 TACGATGTGCTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCGCTG 116  
Db 72 TACGACCTCAGCGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTC 131  
Qy 117 CTGTGCGCTTCCCGCAACAGGTCCACTCCT---CGCACATCTCCACCATCGGTGTT 173  
Db 132 CTGATCCAAATCAAAGACGGGGCTTCTGTCCGGAACCTTCATAGCCACCGTCGGCATA 191  
Qy 174 GACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAAGTGCAGATACAGATCTGGGAC 233  
Db 192 GACTTCAGGAACAAGGTGGTGACTGTGTGATGGCGTGAGAGTGAAGCTGCAGATCTGGAC 251  
Qy 234 ACTCAGGCGAGGAGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCCGAGGG 293  
Db 252 ACCGCTGGCAGGAACGGTTCGGAAGCGTCAACCATGCTTATTACAGAGATGCTCAGGCC 311  
Qy 294 ATATTTTGGTCTATGACTTAGCAGCGGCGCTTACCAGCACATCATGAAGTGGTC 353  
Db 312 TTGCTTCTGTGTATGATCATCCAAACAATCTTCTTCGACAACATCAGGCGCTGGCTC 371  
Qy 354 AGTGACGTGGATGATAGCAGCACAGAGGCGTCCAGAGATGCTTATTGGGAATAAGGCT 413  
Db 372 ACTGAGATTCATGATGATGATGCCAGAGGACGTGTGTGATGCTGTAGGCAACAAGCG 431  
Qy 414 GATGAGAGCAGAAACCGCAGGTGGGAGAGAGAGCAAGGCGAGCAGTGGGGAAGAGTAT 473  
Db 432 GATATGAGCAGGAAGAGTATCGGTCCGAAGACGAGAGACCTTTGGCCAGGGAGTAC 491  
Qy 474 GGCATGACTTCTATGAACAAGTGCCTGACCAACCTCAA 514  
Db 492 GGTGTTCCCTTCTGGAGACAGCGCCCAAGACTGGCATGAA 532

## RESULT 9

US-09-620-312D-959  
; Sequence 959, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 959\_  
; LENGTH: 1106  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (124)..(699)  
; US-09-620-312D-959

Query Match 4.2%; Score 136.2; DB 4; Length 1106;  
Best Local Similarity 57.5%; Pred. No. 6.5e-26;  
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

Qy 57 TACGATGTGCTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCGCTG 116  
Db 103 TACGACCTCAGCGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTC 162  
Qy 117 CTGTGCGCTTCCCGCAACAGGTCCACTCCT---CGCACATCTCCACCATCGGTGTT 173  
Db 163 CTGATCCAAATCAAAGACGGGGCTTCTGTCCGGAACCTTCATAGCCACCGTCGGCATA 222  
Qy 174 GACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAAGTGCAGATACAGATCTGGGAC 233  
Db 223 GACTTCAGGAACAAGGTGGTGACTGTGTGATGGCGTGAGAGTGAAGCTGCAGATCTGGGAC 282  
Qy 234 ACTGAGGCGAGAGAGATACAGACCATCAAAAGCAGTACTATCGGCGGGCCCGAGGG 293  
Db 283 ACCGCTGGCAGGAACGGTTCGGAAGCGTCCAGACGCTCACCATGCTTATTACAGAGATGCTCAGGCC 342  
Qy 294 ATATTTTGGTCTATGACTTAGCAGCGGCGCTTACCAGCACATCATGAAGTGGGTC 353  
Db 343 TTGCTTCTGTGTATGATCATCCAAACAATCTTCTTCGACAACATCAGGCGCTGGCTC 402  
Qy 354 AGTGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 413  
Db 403 ACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462  
Qy 414 GATGAGAGCAGAAACCGGAGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATAT 473  
Db 463 GATATGAGCAGCAAGAGAGTATCGGTTCGGAAGACGAGAGAGACCTTTGGCCAGGGAGTAC 522  
Qy 474 GGCATGACTTCTATGAACAAGTGCCTGACCAACCTCAA 514  
Db 523 GGTGTTCCCTTCTGGAGACAGCGCCCAAGACTGGCATGAA 563

## RESULT 10

US-09-484-970B-142  
; Sequence 142, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkmut, Wayne  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; CURRENT FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 142  
; LENGTH: 2612  
; TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6426186 412477.1CB1  
US-09-484-970B-142

Query Match  
Best Local Similarity 4.2%; Score 136.2; DB 4; Length 2612;  
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;

QY 57 TACGATGCTGTTCCGGCTCTCTGATCGGGACCTCGGGTGGCGAAGACTCGCTG 116  
DB 91 TACGACCTCAGCGGCAAGGTGATCTCTGGAGACACAGCGCTCGGCAAAACATGTTT 150  
QY 117 CTGTGCGCTTACCGCAACAGTTCACCTCT---CGCACATCTCCACCATCGGTGTT 173  
DB 151 CTGATCCANTCAAGACGGGCTTCTCTCGGAACCTTCATAGCCACCGTCGCATA 210  
QY 174 GACTTTAAGATGAAGCATAGAGGTAGACGGCATCAAGTGGCGATACAGATCTGGAC 233  
DB 211 GACTTCAGGAACAAGGTGGTGAATGCTGGTGGAGTGAAGCTGCAGATCTGGAC 270  
QY 234 ACTGCGGCGAGGAGATACCAACCATCAAAAGCAGTACTATCGCGGGCCAGGG 293  
DB 271 ACCGCTGGCGAGGAACGGTCCGAAGCGTCAACCATGCTTTATACAGAGATGCTCAGGCC 330  
QY 294 ATATTTTGGTCTATGACATTAGCAGGAGCGCTCTTACAGCACATCATGAAGTGGTC 353  
DB 331 TTGCTTCTGTTATGATCATCACCAATCTCTTTTGGACATCAGGCGCTGCTC 390  
QY 354 AGTGACGTGGATGATGACGACCAAGAGCGCTCCAGAGATCTTTATTTGGGAATPAGGCT 413  
DB 391 ACTGAGATTCATGATGATGCGCCAGAGGACGCTGTGATCATGCTGTAGGCAACAGGGC 450  
QY 414 GATGAGAGCAAAACGCGAGTGGAGAGAGACGAGCAGAGCTGGCGAAGAGTAT 473  
DB 451 GATATGAGCAGCAAGAGTATCGCTTCCGAAGACGAGAGACCTTTGGCGAGGAGTAC 510  
QY 474 GGCATGACTTCTATGAACAAAGTGCCTGACCAACCTCAA 514  
DB 511 GGTGTTCCCTCTCTGGAGACGAGCGCCAAAGACTGGCATGAA 551

## RESULT 11

US-09-620-312D-646  
Sequence 646, Application US/09620312D  
Patent No. 6569662  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yunqing  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghast  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pt\_FL\_genes Version 1.0  
SEQ ID NO 646  
LENGTH: 1069  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (156)...(839)  
US-09-620-312D-646

Query Match  
Best Local Similarity 4.1%; Score 133.4; DB 4; Length 1069;  
Matches 266; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 52 AGCAGTACGATGCTGCTCCGGCTGCTGATCGGGGACTCGGGTGGCGAAGACT 111  
DB 229 AGAATTTGACTACATGTTCAAAATTACTCATCGGCAATAGCAGTGTGGGAAACAT 288  
QY 112 GCCTGCTGTGCGCTTCCACGCAACAGTTCACCTCTCGCACATCTCCACCATCGGTG 171  
DB 289 CTTTTCTATTCGGTATGAGATGATCTCTTACATCTGCTGAGCAGTTCGGA 348  
QY 172 TTGACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAGTGGCGATACAGATCTGGG 231  
DB 349 TCGATTTCAAAGTAAAACTGTATTCAAAATGAAAAGAGAAATCAAGCTTCAGATTTGGG 408  
QY 232 ACATGCGGCGAGGAGATACAGACCATCAAAAGCAGTACTATCGGGGGCCAGG 291  
DB 409 ACAGCAGCGCGGAGGAAAGATACAGGACTATACACAGCTTATTCGTGGAGCCATGG 468  
QY 292 GGATATTTTGGTCTATGACATTAGCAGCGCGCTCTTACGACCATCATGAAGTGGG 351  
DB 469 GCTTTATTTAATGATGATGACATTTACAATGAGAAATCCTTCAATGAGTACAGATTTGGT 528  
QY 352 TCAGTACGCTGGATGATGACGACCAAGAGCGCTCCAGAGATCTTTATTTGGGAATAGG 411  
DB 529 CAATCAATCAAAACATATCTTTGGGACAAATGCCAAGTTATTCGTGGGAAACAGT 588  
QY 412 CTGATGAGGAGCAGAAACGCGAGGTGGAGAGAGCAAGGCGAGCAGCTGGCGAAGGAT 471  
DB 589 GTGACATGGAAGACGAGCGGCTCATCTCAATGAGCGAGGTCAACATTTAGGAGAACAGC 648  
QY 472 ATGGCATGGACTTCTATGAACAAAGTGCCTGACCAACCTCAACATTTAAAGAGCTTCA 531  
DB 649 TTGGTTTGGATTTTGTGAAACAAAGTCCCAAGGCAACATTAATGTCAGCAGACATTTG 708  
QY 532 CGCGTCT 538  
DB 709 AGCGCT 715

## RESULT 12

US-09-016-434-1422  
Sequence 1422, Application US/09016434  
Patent No. 650938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1422:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G550059
; US-09-016-434-1422

Query Match          3.8%; Score 124.2; DB 4; Length 723;
Best Local Similarity 53.4%; Pred. No. 7.5e-23;
Matches 261; Conservative 0; Mismatches 228; Indels 0; Gaps 0;

QY 55 AGTAGAGTGTCTGTCGGCTGCTGTCGAGTGGGAGTCTCGGGGTGGGCAAGACCTGCC 114
DB |||
QY 70 AATATGATTATTATTCAAGTTACTTCTGATTGGCGACTCAGGGGTTGGAAAGCTTGGCC 129
DB |||
QY 115 TGCTGTGCGCTTCACCGACCAAGAGTTCCACTCTCTCGCATCTCCACCATCGGTGTG 174
DB |||
QY 130 TCTTCTTAGTTTGCAGATGATATATACAGAAAGCTACATCAGACAAATTGGTGTGG 189
QY 175 ACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAAGTCCGGATACAGATCTGGGACA 234
DB |||
QY 190 ATTTCAAAATAAGAACTATAGAGTTAGACGGGAAAAACAATCAAGCTTCANAATATGGACA 249
QY 235 CTGACGGCCAGAGAGATACAGACCATCACAAGACGATCTATCGGGGGCCCGAGGGA 294
DB |||
QY 250 CAGCAGGCCAGAAAGATTTGGAACAATCACCCTCCAGTTATTACAGAGAGCCCATGGCA 309
QY 295 TATTTTGTGTCTATGACATTAGCAGCGGCTCTTACCAGCATCATGAAAGTGGGTCA 354
DB |||
QY 310 TCATAGTTGTGTATGATGTGACAGATCAGGAGTCTTCAATAATGTTAAACAGTGGCTGC 369
QY 355 GTGACGTGGATGAGTACGCCACAGAAAGCGGTCCAGAGATCCTTATTGGGAAATAGGCTG 414
DB |||
QY 370 AGGAAATAGATCGTTATGTCAGTGAATGTCAACAATTTGTTGTAGGGAACAATGTG 429
QY 415 ATGACGACGAGAAACCGCAGAGTGGGAAGACGACGCGCAGCTGGCGAGGAGTATG 474
DB |||
QY 430 ATCTGACCACAAAGAAAGTAGTAGACTACACAACAGCGAAAGAAATTTCTGATTCCTTGT 489
QY 475 GCATGGACTTCTATGAACACAGTGTGCTGCACCACTCAACATTAAGAGTCATTTCAGCG 534
DB |||
QY 490 GAAATTCGGTTTTTGGAAACAGTGTGTAAGNATGCAACGACGTAAGACAGTCTTTCATGA 549
QY 535 GTCTGACAG 543
DB |||
QY 550 CGATGGCAG 558
DB |||

RESULT 13
US-09-620-312D-828
; Sequence 828, Application US/09620312D
; Patent No. 6569662
;
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Aeundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 828
; LENGTH: 1102
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (234)..(875)
; US-09-620-312D-828

Query Match          3.8%; Score 123.6; DB 4; Length 1102;
Best Local Similarity 51.7%; Pred. No. 1.4e-22;
Matches 307; Conservative 0; Mismatches 284; Indels 3; Gaps 1;

QY 12 CGCCCGCAGTTCGCCGCCCGCTGGCCCCAGTGCATCGCGAAGCAGTACGATGTGCTGTTTC 71
DB |||
QY 213 CACCTCCATCGCGAGCCAAAGATGGGAATGGAACCTCAGGAAGATTATATACTTTGTCTTC 272
DB |||
QY 72 CGCGTGTCTGTATCGGGGACTCCGGGTGGGCAAGACCTGCTGCTGTGCCGCTTCACC 131
DB |||
QY 273 AAGGTGGTGTGATCGGCGAATCAGGTGTGGGAAGACCAATCTACTCTCCGATTCAG 332
QY 132 GACAAACAGTTCACATCTCTCGCACATCTCCACCATCGGTGTGACTTTAAGATGAAGACC 191
DB |||
QY 333 CGCAATGAGTTACGCCACAGACGCGCACCATCGGGTTGAGTTCTCCACCGCACT 392
QY 192 ATAGAGGTAGACGGGCATCAAAGTGCAGATACAGATCTGGGACACTGCGAGGCGAGGAGA 251
DB |||
QY 393 GTGATGTGGGCACCGCTGCTGTCAAGGCTCAGATCTGGGACACAGCTGGCGCTGGAGCGG 452
QY 252 TACCAGACATCAAAAGCAGTACTATCGGGGGCCCGAGGGGATATTTTGTGCTATGAC 311
DB |||
QY 453 TACGAGCCATCACCTCGGGGTACTATCTGGGTGTCAGTGGGGGCCCTCTCTGGTTTGAC 512
QY 312 ATTAGCAGCAGCGCTCTTACCAGCATCATGAAGTGGGTCAAGTGAAGTGAAGTAC 371
DB |||
QY 513 CTAACCAAGCACAGACCTATGCTGTGTGGAGCGATGGCTGAAGGAGCTCTATGACCAT 572
QY 372 GCACCAAGAGCGGTCCAGAAAGATCCTTATTGGGAATGAAGCTGATGAGGAGCAGAAACGG 431
DB |||
QY 573 GCTGAAGCCACGATCGTCTGTCATGCTCGTGGGTAAACAAAAGTGAACCTCAGCAGGCCCGG 632
QY 432 CAGGTGGGAAGAGAGCAAGGCGACGCTGGCGAAGGAGTATGGCATGGACTTCTATGAA 491
DB |||
QY 633 GAAGTGCCCACTGAGGAGGCCCGAATGTTTCGCTGNAACATGAGACTGCTCTTCTCTGAG 692
DB |||
```

Qy 492 ACAAGTGCCTGCACCAACCTCAACATTAA---AGAGTCATTACGCGTCTGACAGAGCTG 548  
Db 693 ACCTCAGCCCTGGACTCTACCAATGTTGAGCTAGCTTTGAGACTGTCTTGAAGAAATC 752  
Qy 549 GTGCTGAGGCCCTTAGAAGAGAGCTGGAAGCCTCCGGATGCGTGCAGCAAT 602  
Db 753 TTTCCGAAGGTGTCCAGCAGAGACAGAACAGCATCCGGACCAATGCCATCACT 806

RESULT 14  
US-09-566-921-7  
; Sequence 7, Application US/09566921  
; Patent No. 6682888  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.  
; APPLICANT: Edwards, Carla M.  
; FILE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE  
; TITLE OF INVENTION: PA-0024 US  
; CURRENT APPLICATION NUMBER: US/09/566,921  
; CURRENT FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: PERL Program  
; SEQ ID NO 7  
; LENGTH: 8137  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6682888 411474.10  
; NAME/KEY: unsure  
; LOCATION: 3488-3788  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-566-921-7

Query Match 3.8%; Score 122.6; DB 4; Length 8137;  
Best Local Similarity 53.2%; Pred. No. 7.7e-22;  
Matches 260; Conservative 0; Mismatches 229; Indels 0; Gaps 0;  
Qy 55 AGTAGATGCTGTTCCGGCTGCTGTCGCGGACTCGGGGACTCGGGGTGGGCAAGACTGCC 114  
Db 222 AATATGATTATTATTCAAGTTACTTCTGATGGCGACTCAGGGGTTGGAAAGTCTTGCC 281  
Qy 115 TGCTGTGCGCTTCACGCACAAACAGTTCCTCCTCGCACATCTCCACCATCGTGTTG 174  
Db 282 TCTCTTAGTGTGCGATGATACATATACAGAAAGCTACATGACAAATGGTGTGG 341  
Qy 175 ACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAAAGTCGGGATACAGATCTGGGACA 234  
Db 342 ATTTCAAAATAAGAACTATAGATTAGACGGGAAACCAATCAAGTTCAATATGGGACA 401  
Qy 235 CTGAGGCGAGGAGATACAGACCATCAAAAGCAGTACTATCGCGGGGCCCGGAGGA 294  
Db 402 CAGCAGGCGAGGAAAGATTTCGAACAATCACCTCCAGTTATTACAGAGAGCCCATGCA 461  
Qy 295 TATTTTGGTCTGTACATTAGCAGCGCTCTTACGACACATCATGAAGTGGGTCA 354  
Db 462 TCATAGTTGTGTATGTGACAGATCAGGAGTCTCTCAATAATGTTAAACAGTGGTGC 521  
Qy 355 GTGACGTGGATGAGTACGACAGAGCGGTCCAGAGAGATCTTTATTTGGGAATAAGGCTG 414  
Db 522 AGGAAATAGATCGTTATCCAGTGAATATGTCAACAATTTGTGTAGGAAACAAATGTG 581  
Qy 415 ATGAGGACAGAAACGGCAGGTGGGAGAGAGCAGGCGCAGCTGGCGAAGGATG 474  
Db 582 ATCTGACCAACAAGAAAGTAGTAGACTACACAACAGCGAAGGAATTTGCTGATTCCTT 641  
Qy 475 GCATGGACTTCTATGAACAGTGCCTGCCCAACCTCAACATTTAAAGAGTCATTCAGC 534  
Db 642 GAAATCCGCTTTTGGAAACCAAGTGTGAAGATGCAACGAATGTAGAACAGTCTTTTCATGA 701  
Qy 535 GTCTGACAG 543

Db 702 CGATGGCAG 710  
RESULT 15  
US-08-773-423-4  
; Sequence 4, Application US/08773423  
; Patent No. 5869291  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Bandman, Olga  
; TITLE OF INVENTION: NOVEL RAB PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/773,423  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0183 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 847 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Consensus  
; CLONE: Consensus  
US-08-773-423-4

Query Match 3.7%; Score 120.6; DB 2; Length 847;  
Best Local Similarity 51.5%; Pred. No. 7.3e-22;  
Matches 301; Conservative 0; Mismatches 280; Indels 3; Gaps 1;  
Qy 22 TCCCGGCGCGCTCGCCCGCTCATGCGAAGCAGTAGTGTCTGTTCCGGCTGCTGC 81  
Db 60 TCGCGANCAAGATGGGGAATGGAATCTGAGGAGATTAATCTTGTCTTCAAGTGTGTC 119  
Qy 82 TGATCGGGGAGCTCCGGGGTGGCAAGACCTGCTGCTGTGCGGCTTCCCGCAACAGT 141  
Db 120 TGATCGCGGAATCAGGTGTGGGGAAGCAATCTACTCTCCGATTCACCGGCAATGAT 179  
Qy 142 TCCACTCTCGCATCTCCACCATCGGTGTGTGACTTTAAGATGAAGACCATAGAGTAG 201  
Db 180 TCAGCCACGACAGCGCAGCCACCATCGGGGTGTGAGTTCTCCACCGGCACTGTGTGTTGG 239  
Qy 202 ACGGCATCAAGTCCGGATACAGATCTGGGACACTGCGAGGCGAGGAGATACAGACCA 261  
Db 240 GCACCGTGTGTCAAGGCTCAGATCTGGGACACAGCTGGGCTGGAGGGTACCGAGCCA 299  
Qy 262 TCACAAAGCAGTACTATATCGGGGGCCCGAGGGGATATTTTGGTCTATGACATTAGCAGC 321



Db	300	TCACCTCGGCGTACTATCGTGGTGCAGTGGGGCCCTCCTGGTGTGACCTAACCAAGC	359
Qy	322	AGCGCTTTACCAAGCAATCATGAAGTGGGTGAGTGCAGTGGATGAGTACGACCAAG	381
Db	360	ACCAGACCTATGCTGTGGTGGAGCGATGGCTGAAGGAGCTCTATGACCATGCTGAAGCCA	419
Qy	382	CGGTCCAGAGATCCTTATTGGGATTAAGGCTGATGAGGAGCAGAACGGCAGGTGGGA	441
Db	420	CGATCGTGCATGCTCGTGGGTAAACAAAAGTGACCTCAGCCAGGGCCGGGAAGTGCCCA	479
Qy	442	GAGAGCAAGGCGACAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAACAAGTGCCT	501
Db	480	CTGAGGAGGCCGGAATGTTGCTGAAAACAATGGACTGCTTCTCTGGAGACCTCAGCCC	539
Qy	502	GCACCAACCTCAACATTAA---AGAGTCAATTCACGCGTCTGACAGAGCTGGTGTGAGG	558
Db	540	TGGACTCTACCAATGTTGAGCTAGCCTTTGAGACTGTCTCTGAAGAAATCTTTGCGAAGG	599
Qy	559	CCCATAGGAGGAGCTGGAAGGCGCTCCGGATGCGTGCCGCAAT	602
Db	600	TGTCCAAGCAGAGACAGAACAGCATCCGGACCAATGCCATCACT	643

Search completed: March 21, 2004, 02:20:55  
Job time : 231 secs



Db 181 AGATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACACTGCAG 240  
Qy  
Qy 241 GGCAGGAGAGATACACAGCATCAAAAGCAGTACTATCGCGCGGCCACCGGATATTTT 300  
Db 241 GGCAGGAGAGATACACAGCATCAAAAGCAGTACTATCGCGCGGCCACCGGATATTTT 300  
Qy 301 TGGTCTATGACATTAAGCAGGAGCGCTCTTACACAGCACATCATGAAGTGGGTGAGTGAAG 360  
Db 301 TGGTCTATGACATTAAGCAGGAGCGCTCTTACACAGCACATCATGAAGTGGGTGAGTGAAG 360  
Qy 361 TGGATGATGACGACGAGAGAGCGCTCAGAAAGTCTTATTTGGGAATAAGGCTGATGAGG 420  
Db 361 TGGATGATGACGACGAGAGAGCGCTCAGAAAGTCTTATTTGGGAATAAGGCTGATGAGG 420  
Qy 421 AGCAGAAACGCGAGTGGGAAGAGAGCAAGGCGAGCAGCTGGCGAAGAGTATGGCATGG 480  
Db 421 AGCAGAAACGCGAGTGGGAAGAGAGCAAGGCGAGCAGCTGGCGAAGAGTATGGCATGG 480  
Qy 481 ACTTCTATGAACAAAGTGCCTGCAACCACTCAACATTTAAAGAGTCAATTCACGCGTCTGA 540  
Db 481 ACTTCTATGAACAAAGTGCCTGCAACCACTCAACATTTAAAGAGTCAATTCACGCGTCTGA 540  
Qy 541 CAGAGCTGGTCTGACGCGCCATAGGAAGAGCTGGAAAGGCTCCGGATGGGTGCCAGCA 600  
Db 541 CAGAGCTGGTCTGACGCGCCATAGGAAGAGCTGGAAAGGCTCCGGATGGGTGCCAGCA 600  
Qy 601 ATGAGTTGGCACTGGCAGAGCTGAGGAGGAGGAGGCAACCCGAGGCGCCACGCAACT 660  
Db 601 ATGAGTTGGCACTGGCAGAGCTGAGGAGGAGGAGGCAACCCGAGGCGCCACGCAACT 660  
Qy 661 CTTTCGAAAACTGCTGTGTGTGAGTCTGTGTGGGCAACCCACACAGCACACCCCTCTTTC 720  
Db 661 CTTTCGAAAACTGCTGTGTGTGAGTCTGTGTGGGCAACCCACACAGCACACCCCTCTTTC 720  
Qy 721 CTCAGGAGGCGCGTGGGAGACAGGAGGAGCGGGCTTTGGCCCTGTGCTGTCTCTCGT 780  
Db 721 CTCAGGAGGCGCGTGGGAGACAGGAGGAGCGGGCTTTGGCCCTGTGCTGTCTCTCGT 780  
Qy 781 GTGATGACCTTATGAGTATCAGTAGCCACTACTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 GTGATGACCTTATGAGTATCAGTAGCCACTACTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Qy 841 CTGCTGTCTCATCAAGCAGCGCCCTGTGCCAGCGCGTCCACCTTGGAGTGGTCTTCTTCA 900  
Db 841 CTGCTGTCTCATCAAGCAGCGCCCTGTGCCAGCGCGTCCACCTTGGAGTGGTCTTCTTCA 900  
Qy 901 GCCTGTTTCCCGAGCAGCGCTGTGCCAGCGCGCTGAGTGTGCGCGGAGCACTGTCTC 960  
Db 901 GCCTGTTTCCCGAGCAGCGCTGTGCCAGCGCGCTGAGTGTGCGCGGAGCACTGTCTC 960  
Qy 961 ACCATCCGCAACCCACAGCAACAGCCAGGCGTGGAGTCCAGGCACTTTTCAGCTGCTC 1020  
Db 961 ACCATCCGCAACCCACAGCAACAGCCAGGCGTGGAGTCCAGGCACTTTTCAGCTGCTC 1020  
Qy 1021 CTTTCTCGTGCATCGTGTCTTCTGTGCTTTTTCTCTCTTCCCGCACTTCTCTTCTC 1080  
Db 1021 CTTTCTCGTGCATCGTGTCTTCTGTGCTTTTTCTCTCTTCCCGCACTTCTCTTCTC 1080  
Qy 1081 TGACCCCTCCCTCGGTGCGTTTGGTATCAAGCTCCTCAAAACCCGCTCCCGGTGTGT 1140  
Db 1081 TGAACCCCTCCCTCGGTGCGTTTGGTATCAAGCTCCTCAAAACCCGCTCCCGGTGTGT 1140  
Qy 1141 CTTGCTGTGTCAGTCTGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
Db 1141 CTTGCTGTGTCAGTCTGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
Qy 1201 GCTCTGGGAGGTTCACCTTGGATCCAGGAAGACCTCCACCTCGCTCGTGGGTG 1260  
Db 1201 GCTCTGGGAGGTTCACCTTGGATCCAGGAAGACCTCCACCTCGCTCGTGGGTG 1260  
Qy 1261 GGCCAAAGGCTACAGGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320  
Db 1261 GGCCAAAGGCTACAGGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320

Qy 1321 TGGGCCTGCTCTCCAGTGAACCTTGCAGAAAGTGGAGCATCGAGGTAGGAGGAAACAGCAA 1380  
Db 1321 TGGGCCTGCTCTCCAGTGAACCTTGCAGAAAGTGGAGCATCGAGGTAGGAGGAAACAGCAA 1380  
Qy 1381 CCGGGAGTCTCGAGCCTGGGGCTGCCCTTACCTCTACCCATTTCCCGGACAGAGCTTTG 1440  
Db 1381 CCGGGAGTCTCGAGCCTGGGGCTGCCCTTACCTCTACCCATTTCCCGGACAGAGCTTTG 1440  
Qy 1441 CCTTCTTGGCTGCGCGCTGCTCTTTGGGAACTGAGCTCAGAGGAGGTGCTTCAG 1500  
Db 1441 CCTTCTTGGCTGCGCGCTGCTCTTTGGGAACTGAGCTCAGAGGAGGTGCTTCAG 1500  
Qy 1501 AGAAGGAAACAAATGAGGGGTGGCAGGGATAAAAAGTCACTCCATTTCTTCTACCTCCCA 1560  
Db 1501 AGAAGGAAACAAATGAGGGGTGGCAGGGATAAAAAGTCACTCCATTTCTTCTACCTCCCA 1560  
Qy 1561 TGAGCATGAACAAATTTCTTCCACTGGCTCCCAATTTTAAAGTGGACCAAGGC 1620  
Db 1561 TGAGCATGAACAAATTTCTTCCACTGGCTCCCAATTTTAAAGTGGACCAAGGC 1620  
Qy 1621 CTGTGGGTACTCCAGGGCAAGGAGAGCCCTGGGTGAGTCACTGTCAAGGCAACCAT 1680  
Db 1621 CTGTGGGTACTCCAGGGCAAGGAGAGCCCTGGGTGAGTCACTGTCAAGGCAACCAT 1680  
Qy 1681 GCATCTCCACAAAGGGAGCATTTTGGAAATGAAGGACTAGTCTCTATGTATCAGGTTAAGA 1740  
Db 1681 GCATCTCCACAAAGGGAGCATTTTGGAAATGAAGGACTAGTCTCTATGTATCAGGTTAAGA 1740  
Qy 1741 GCAAGGAGAGCTGGCCAGGGAACAGAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800  
Db 1741 GCAAGGAGAGCTGGCCAGGGAACAGAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800  
Qy 1801 GGGCCTCTTAGGCGCCCATCTTCCATTTCTTAGGTAAAGAGCATTTCTCAGACTCCCA 1860  
Db 1801 GGGCCTCTTAGGCGCCCATCTTCCATTTCTTAGGTAAAGAGCATTTCTCAGACTCCCA 1860  
Qy 1861 GGGCCTCTTAGGCGCCCATCTTCCATTTCTTAGGTAAAGAGCATTTCTCAGACTCCCA 1920  
Db 1861 GGGCCTCTTAGGCGCCCATCTTCCATTTCTTAGGTAAAGAGCATTTCTCAGACTCCCA 1920  
Qy 1921 GAGAAGGCAAGACTTTCATGGGAAGAGAAAGAGGCGCTGGGTAGAAACCGTTGGT 1980  
Db 1921 GAGAAGGCAAGACTTTCATGGGAAGAGAAAGAGGCGCTGGGTAGAAACCGTTGGT 1980  
Qy 1981 CTGTCTCTTTGGCTTTAAGACAAAGCGCTCATCTTGGCTCTTACCTCTGTAGTGGCT 2040  
Db 1981 CTGTCTCTTTGGCTTTAAGACAAAGCGCTCATCTTGGCTCTTACCTCTGTAGTGGCT 2040  
Qy 2041 GAGGTTTGCACACACTGTGGCTACAGGTGGAGGAAAGAGACTCTTCTTCTCCAGAG 2100  
Db 2041 GAGGTTTGCACACACTGTGGCTACAGGTGGAGGAAAGAGACTCTTCTTCTCCAGAG 2100  
Qy 2101 TGCTATGTTGAGAGGTTCTTTAAACCCCATATGGCCCAAGTAGTCTGTAGAGGCCC 2160  
Db 2101 TGCTATGTTGAGAGGTTCTTTAAACCCCATATGGCCCAAGTAGTCTGTAGAGGCCC 2160  
Qy 2161 TTTTAAAGACGGAACAAGTATTTTACAGTTTCTTCTGGGGTCTTCTGGCCACCGTCCCAAG 2220  
Db 2161 TTTTAAAGACGGAACAAGTATTTTACAGTTTCTTCTGGGGTCTTCTGGCCACCGTCCCAAG 2220  
Qy 2221 TGGCGAGGCTTAGGAAGAGGTCATTTCTTAAAGCCACACATTAGTCTGCTGCTGCTG 2280  
Db 2221 TGGCGAGGCTTAGGAAGAGGTCATTTCTTAAAGCCACACATTAGTCTGCTGCTGCTG 2280  
Qy 2281 CAGCAAAAACAAAGAACTGGGTGTGAGTATTCATCAACTAAGAACCAAAATCCAGGGCA 2340  
Db 2281 CAGCAAAAACAAAGAACTGGGTGTGAGTATTCATCAACTAAGAACCAAAATCCAGGGCA 2340  
Qy 2341 CTCATATGTGAAGATGAAGACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400  
Db 2341 CTCATATGTGAAGATGAAGACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400

```
Qy 2401 ACCATCAAACTTTCTCTCTGACTTACAAACCAAGGAAAAACAGCAGAGAGGGTGGCTCA 2460
Db 2401 ACCATCAAACTTTCTCTCTGACTTACAAACCAAGGAAAAACAGCAGAGAGGGTGGCTCA 2460
Qy 2461 GGACTTAGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGAGGAGACGAGAGCTTGTA 2520
Db 2461 GGACTTAGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGAGGAGACGAGAGCTTGTA 2520
Qy 2521 ATCACTGGCTAATCAGAAAAAGGACAGACTAACTCTAGGATGAAGCTGTGACTAGGCTGG 2580
Db 2521 ATCACTGGCTAATCAGAAAAAGGACAGACTAACTCTAGGATGAAGCTGTGACTAGGCTGG 2580
Qy 2581 AGTTGCTTCTTGAAGTGGACCTCTTGGGTATCAAGACCTATGCCCCATACACTGGG 2640
Db 2581 AGTTGCTTCTTGAAGTGGACCTCTTGGGTATCAAGACCTATGCCCCATACACTGGG 2640
Qy 2641 GCTTAGGGAAGTAGTGATGCCAGGCCCTCAAGCTCTGCTTTCAGCCAGGACTTGAAGATT 2700
Db 2641 GCTTAGGGAAGTAGTGATGCCAGGCCCTCAAGCTCTGCTTTCAGCCAGGACTTGAAGATT 2700
Qy 2701 ATATTGGCAGTGGCTCCAACTCTGTGACCAAGTATTTTCCCTGAAAGATCAGGCA 2760
Db 2701 ATATTGGCAGTGGCTCCAACTCTGTGACCAAGTATTTTCCCTGAAAGATCAGGCA 2760
Qy 2761 GGGTGCCATTCAATTGCTTTCTCTCTAGCCCCCTCAGGAAAGAGACTATATTGTGAC 2820
Db 2761 GGGTGCCATTCAATTGCTTTCTCTCTAGCCCCCTCAGGAAAGAGACTATATTGTGAC 2820
Qy 2821 TGTACCTTAGGGTCTGGAAAGGAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 2880
Db 2821 TGTACCTTAGGGTCTGGAAAGGAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 2880
Qy 2881 CTATCCACAAGGGCCATGACTGGGAAAGATATGGGACAGAGAGATTTGGGATTTTA 2940
Db 2881 CTATCCACAAGGGCCATGACTGGGAAAGATATGGGACAGAGAGATTTGGGATTTTA 2940
Qy 2941 GGGTGCACTACGCTCACTTAACTTTTGTGGCCCTGGGGCATGTCTTGAGGCCCAAGAC 3000
Db 2941 GGGTGCACTACGCTCACTTAACTTTTGTGGCCCTGGGGCATGTCTTGAGGCCCAAGAC 3000
Qy 3001 TGTAAAGCAGGCTCTGTGGCTGTACTCGTACCACTCTGACCTCTGCTGTCTGTGAG 3060
Db 3001 TGTAAAGCAGGCTCTGTGGCTGTACTCGTACCACTCTGACCTCTGCTGTCTGTGAG 3060
Qy 3061 ACTCCATCCAGCCAGGACGACCGCTCTGAGCCCTCCACTATCTCCCTGTGAGGG 3120
Db 3061 ACTCCATCCAGCCAGGACGACCGCTCTGAGCCCTCCACTATCTCCCTGTGAGGG 3120
Qy 3121 GTGAACCTTCTGTGTCTCGGGTCCATATATGAATTGTGAGCAGGGTTCTATTT 3180
Db 3121 GTGAACCTTCTGTGTCTCGGGTCCATATATGAATTGTGAGCAGGGTTCTATTT 3180
Qy 3181 TTAACACAGATGTTTACAAAATAAAGATTATTTCAACACCAACCAAAAAA 3240
Db 3181 TTAACACAGATGTTTACAAAATAAAGATTATTTCAACACCAACCAAAAAA 3240
Qy 3241 AAAAAAAAAAAAAA 3257
Db 3241 AAAAAAAAAAAAAA 3257
```

RESULT 2

```
US-09-817-198A-3
; Sequence 3, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: PROTEINS, AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
```

```
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28770
; TYPE: DNA
; ORGANISM: Human
US-09-817-198A-3
```

```
Query Match 82.8%; Score 2696.6; DB 9; Length 28770;
Beat Local Similarity 98.8%; Pred. No. 0;
Matches 2717; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
```

```
Qy 473 TGGCATGGACTTCTATGAACAAAGTGCCTGCACCAACTCAACATTTAAAGATGATTTAC 532
Db 25562 TGCAGGGGAGAGGAGGAGACACTGAGCTTACCTGTGCTTTGGTTTCCAGTCAATTCAC 25621
Qy 533 GCGTCTCAGACAGCTGTGTGTCAGAGGCCCATAGGAAGAGCTGGAAGCCCTCCGATGCG 592
Db 25622 GCGTCTCAGACAGCTGTGTGTCAGAGGCCCATAGGAAGAGCTGGAAGCCCTCCGATGCG 25691
Qy 593 TGCAGCAATGAGTTGGCACTTGGCAGAGCTGGAGGAGGAGGAGGCAAAACCCGAGGGCCC 652
Db 25682 TGCAGCAATGAGTTGGCACTTGGCAGAGCTGGAGGAGGAGGAGGCAAAACCCGAGGGCCC 25741
Qy 653 AGCGAACTCTTCGAAAACTGTGCTGAGTCTGTGTGGGGCACCCCAACAGACACC 712
Db 25742 AGCGAACTCTTCGAAAACTGTGCTGAGTCTGTGTGGGGCACCCCAACAGACACC 25801
Qy 713 CCTCTTCCCTCAGAGAGCCGTGGGACAGACAGGAGGAGCCGGGCTTTGCCCTGCTGCTGT 772
Db 25802 CCTCTTCCCTCAGAGAGCCGTGGGACAGACAGGAGGAGCCGGGCTTTGCCCTGCTGCTGT 25861
Qy 773 CCTCTGCTGTGATCACCCTATTGAGTATCAGTAGCCACTACTCTCCCTGCTGGCCCTGA 832
Db 25862 CCTCTGCTGTGATCACCCTATTGAGTATCAGTAGCCACTACTCTCCCTGCTGGCCCTGA 25921
Qy 833 GAGCGGCTCTGCTGTCTATCTCAAGCAGCCCTGTGCCAGCCCTGCCACCTGGAGTGGT 892
Db 25922 GAGCGGCTCTGCTGTCTATCTCAAGCAGCCCTGTGCCAGCCCTGCCACCTGGAGTGGT 25981
Qy 893 CTCTCTCAGCCCTGTTTCCAGCCACAGAGCCCTGTACGACCCCAACGATGTCCCGCAAGC 952
Db 25982 CTCTCTCAGCCCTGTTTCCAGCCACAGAGCCCTGTACGACCCCAACGATGTCCCGCAAGC 26041
Qy 953 ACTGCTCAACCATCCGACCCACAGACACAGCAGGGGTGGAGTCCAGGCCACTTTC 1012
Db 26042 ACTGCTCAACCATCCGACCCACAGACACAGCAGGGGTGGAGTCCAGGCCACTTTC 26101
Qy 1013 AGCTGCTCTCTTCTCCGTGCACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1072
Db 26102 AGCTGCTCTCTTCTCCGTGCACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26161
Qy 1073 TCTTTCTGACCCCTCCCTCCCGTGGGTTTGTATCAAGCTCTCTCAAAACCCCGTCCC 1132
Db 26162 TCTTTCTGACCCCTCCCTCCCGTGGGTTTGTATCAAGCTCTCTCAAAACCCCGTCCC 26221
Qy 1133 CCGTGTGCTCTGTGTGACAGCTCGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1192
Db 26222 CCGTGTGCTCTGTGTGACAGCTCGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26281
Qy 1193 GGACCCAGGCTCTGTTGGGAGGTTCCACCTTTGGATCCAGGAAGAACCTTCCACCTCGCT 1252
Db 26282 GGACCCAGGCTCTGTTGGGAGGTTCCACCTTTGGATCCAGGAAGAACCTTCCACCTCGCT 26341
Qy 1253 CGTGGGTGGGCAAAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1312
Db 26342 CGTGGGTGGGCAAAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26401
Qy 1313 ATGTGCCATGGCTGCTCTCCAGTACCTGCGGAAGTGGAGCATCCAGGTAGGAGGA 1372
Db 26402 ATGTGCCATGGCTGCTCTCCAGTACCTGCGGAAGTGGAGCATCCAGGTAGGAGGA 26461
Qy 1373 AACAGCAACCGGGGAGTCTCTGAGCCTGGGGCTGCCCTACCTCTACCTATTTCCCGACCA 1432
```

Db 26462 AACGGCAACAGGGAGTCTCGAGCTTGGGGCTGCCCTACCTCTACCCATTCCTCCCGACCA 26521  
Qy 1433 GAGCTTTGCCCTTGTCTTGGCTGCCCGCTGCCTCTTTTGGGAACTGAGCTCAGAGGCGG 1492  
Db 26522 GAGCTTTGCCCTTGTCTTGGCTGCCCGCTGCCTCTTTTGGGAACTGAGCTCAGAGGCGG 26581  
Qy 1493 TGCTTCAGAGAGGAAACAAATAGAGGGTGGCAGGGATAAAAGTCACCTCCATCTCT 1552  
Db 26582 TGCTTCAGAGAGGAAACAAATAGAGGGTGGCAGGGATAAAAGTCACCTCCATCTCT 26641  
Qy 1553 ACCTCCATGAGATGAAACAAATTTCTCTCCACTGGCTCCCAATTTAAAGATGTGG 1612  
Db 26642 ACCTCCATGAGATGAAACAAATTTCTCTCCACTGGCTCCCAATTTAAAGATGTGG 26701  
Qy 1613 ACCAAGCCCTGTGGGTACTCCAGGGCAAGAGAGCCCTGGGTCACTGACACTGTGAG 1672  
Db 26702 ACCAAGCCCTGTGGGTACTCCAGGGCAAGAGAGCCCTGGGTCACTGACACTGTGAG 26761  
Qy 1673 CCAACCATGCATCCCAAGGGGAGCATTTGGAAATGAAGGACTAGCTCCTATGTATCA 1732  
Db 26762 CCAACCATGCATCCCAAGGGGAGCATTTGGAAATGAAGGACTAGCTCCTATGTATCA 26821  
Qy 1733 GGTTAAGACAAGGAGAGCTGGCCAGGACAGCAGTTTGCACAGAGAGGGGAATGTAG 1792  
Db 26822 GGTTAAGACAAGGAGAGCTGGCCAGGACAGCAGTTTGCACAGAGAGGGGAATGTAG 26881  
Qy 1793 CAACAGCAGGGCTCTAGGCCCATCTTCCATTTCTTAGTAAAGAGCATTTCCCTCA 1852  
Db 26882 CAACAGCAGGGCTCTAGGCCCATCTTCCATTTCTTAGTAAAGAGCATTTCCCTCA 26941  
Qy 1853 GACTCCAGGCGAGGACTAGCCCTAGCCCTTACCAACCAAGGTTCTCTGGGACCCAAA 1912  
Db 26942 GACTCCAGGCGAGGACTAGCCCTAGCCCTTACCAACCAAGGTTCTCTGGGACCCAAA 27001  
Qy 1913 GTTTATGGGAGAGGGCAAGACATTCATTTGGGAAGAGAGAGGAGGCTGGGTAGAAC 1972  
Db 27002 GTTTATGGGAGAGGGCAAGACATTCATTTGGGAAGAGAGAGGAGGCTGGGTAGAAC 27061  
Qy 1973 GCTTGTGTCTTCTTCTTGGCTTTAAGACAAGCGCTCATCTTGGCTCTACTCTCTG 2032  
Db 27062 GCTTGTGTCTTCTTCTTGGCTTTAAGACAAGCGCTCATCTTGGCTCTACTCTCTG 27121  
Qy 2033 ATAGCTTGAGGGTTTGCCAACACACACTGTGGCTACAGAGTGGAGGAGAGGACTCCTTC 2092  
Db 27122 ATAGCTTGAGGGTTTGCCAACACACACTGTGGCTACAGAGTGGAGGAGGACTCCTTC 27181  
Qy 2093 CTCAGAGTGTATGTTTCAAGAACTTTCTTAAACCCCATATGGCCCAAGAGTAGCTCGTA 2152  
Db 27182 CTCAGAGTGTATGTTTCAAGAACTTTCTTAAACCCCATATGGCCCAAGAGTAGCTCGTA 27241  
Qy 2153 GGAGGCCCTTTAAGACGGACAGTAATTTACAGTTCTACTGGGGTTCCTGCCACCG 2212  
Db 27242 GGAGGCCCTTTAAGACGGACAGTAATTTACAGTTCTACTGGGGTTCCTGCCACCG 27301  
Qy 2213 TCCCAAGTGGCGAGGCTAGGAGAGGGTCACTTTAAGCCACACATAGCTGCACTG 2272  
Db 27302 TCCCAAGTGGCGAGGCTAGGAGAGGGTCACTTTAAGCCACACATAGCTGCACTG 27361  
Qy 2273 CGTGGCTGACCCAAAACAAAGAACTGGGTGTGAGTATTCATCACTAAGAACCAAAAT 2332  
Db 27362 CGTGGCTGACCCAAAACAAAGAACTGGGTGTGAGTATTCATCACTAAGAACCAAAAT 27421  
Qy 2333 CCAGGCACTATGTGAAGATGAAGACCTCTACTCTTCTTCTCCCAAGAGAGTG 2392  
Db 27422 CCAGGCACTATGTGAAGATGAAGACCTCTACTCTTCTTCTCCCAAGAGAGTG 27481  
Qy 2393 GGGAAAGAACCATCAAACTTCTCTCTGACTTACCAAAACAGGAAACACAGAGAGG 2452  
Db 27482 GGGAAAGAACCATCAAACTTCTCTCTGACTTACCAAAACAGGAAACACAGAGAGG 27541  
Qy 2453 GTGGCTCAGGACTTAGGACAGGGTATAGCTTAGATGGTGAAGAGCAAGAGAGAGGAG 2512

Db 27542 GTGGCTCAGGACTTAGGACAGGGTATAGCTTAGATGGTGAAGCAAGAGGAGAGGAG 27601  
Qy 2513 AGTTGTAATCACTGGCTAATGAGAAAGAGAGACAGCTAACTCTAGATGAAGCTGTGAC 2572  
Db 27602 AGTTGTAATCACTGGCTAATGAGAAAGAGAGACAGCTAACTCTAGATGAAGCTGTGAC 27661  
Qy 2573 TAGGCTGGAGTTCCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 2632  
Db 27662 TAGGCTGGAGTTCCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 27721  
Qy 2633 ACACCTGGGCTAGGAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAAGGAGGACTT 2692  
Db 27722 ACACCTGGGCTAGGAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAAGGAGGACTT 27781  
Qy 2693 GAGAAGTTATTTGGGCGAGTGGCTCCAATCTGTGGACCAAGTATTTCAAGCTTTCCCTGAAG 2752  
Db 27782 GAGAAGTTATTTGGGCGAGTGGCTCCAATCTGTGGACCAAGTATTTCAAGCTTTCCCTGAAG 27841  
Qy 2753 ATCAGGAGGGTGCCTATTCATTTCTTCTCTAGCCCCCTCAGGAAAGAGGACTAT 2812  
Db 27842 ATCAGGAGGGTGCCTATTCATTTCTTCTCTAGCCCCCTCAGGAAAGAGGACTAT 27901  
Qy 2813 ATTTGTACTGTACCTAGGGTTCTGGAAGGAAACATGGAATCAGGATTTCTATAGACT 2872  
Db 27902 ATTTGTACTGTACCTAGGGTTCTGGAAGGAAACATGGAATCAGGATTTCTATAGACT 27961  
Qy 2873 GATAGGCCCTTATCCAAAGGGCCATGACTGGGAAAGGTATGGGAGCAGAGAGGAAATTG 2932  
Db 27962 GATAGGCCCTTATCCAAAGGGCCATGACTGGGAAAGGTATGGGAGCAGAGAGGAAATTG 28021  
Qy 2933 GGAATTTAGGGTGCAGCTAGCTCACCTTAAACCTTTTGGTGGCTGGGGCATGTCTTGAG 2992  
Db 28022 GGAATTTAGGGTGCAGCTAGCTCACCTTAAACCTTTTGGTGGCTGGGGCATGTCTTGAG 28081  
Qy 2993 GCCCAGACTGTTAAGCAGGCTCTGCTGGCTGTTTACTCGTCAACCTCTGCACCTGCT 3052  
Db 28082 GCCCAGACTGTTAAGCAGGCTCTGCTGGCTGTTTACTCGTCAACCTCTGCACCTGCT 28141  
Qy 3053 GTCTGAGACTCCATCCAGCCCCCAGGACGCCACCTGCTCCTGAGCCTCCACTATCTCCC 3112  
Db 28142 GTCTGAGACTCCATCCAGCCCCCAGGACGCCACCTGCTCCTGAGCCTCCACTATCTCCC 28201  
Qy 3113 TGTGAGGGTGAACCTTGTGTACTGTCTCGGGTCCATATATGAATTTGAGCAGGGTT 3172  
Db 28202 TGTGAGGGTGAACCTTGTGTACTGTCTCGGGTCCATATATGAATTTGAGCAGGGTT 28261  
Qy 3173 CATCTATTTTAAACACAGATGTTTACAAATAAGATTTTCAAAACCC 3223  
Db 28262 CATCTATTTTAAACACAGATGTTTACAAATAAGATTTTCAAAACCC 28312

## RESULT 3

US-10-311-455-2111  
; Sequence 2111, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determini  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311.455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 2111  
; LENGTH: 7924

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2111

Query Match      36.6%; Score 1192.8; DB 14; Length 7924;
Best Local Similarity 75.8%; Pred. No. 0;
Matches 1476; Conservative 0; Mismatches 472; Indels 0; Gaps 0;

Qy 1276 GGTGCTTCTTCTCTCTCCACCCACCTGTCTCCTCATGTGCGCATCGGGCTGCTCCCC 1335
Db 1 GGTGCTTTTTTTTTTTTTTATTTTTATTTTTATGTTTTTATGTTGTTAUGGTTTGTTTTT 60

Qy 1336 AGTGACCTCGAAGTAGGAGCATCGAGGTAGAGGAGAAACAGCAACCGGGAGTCTCGA 1395
Db 61 AGTGATTTGCAAGTAGTGAGTATCGAGTAGGAGGAGAAACGGTAATTAGCGAGTTTCGA 120

Qy 1396 GCCTGGGGCTGCCCTACCTCTACCCATTCCCCGACAGAGCTTTGCCCTTGTGGCTGC 1455
Db 121 GTTTGGGGTGTGTTTTATTTTTATTTATTTTTTCGATTAGAGTTTGTTTTTTGTGTGT 180

Qy 1456 CCGCTCGCTCTTTTGGGAACTGAGCTCAGAGGCAGGTGCTTCAGAGCAAGAAACAAAT 1515
Db 181 TCGTTTGTTTTTTTTGGGAAATGAGTTTAGGTAGGTGTTTTAGAGAGAAATAAAT 240

Qy 1516 GAGGGGTGGCAGGGATAAAAGTCACCTCCAATCTCACTCCCATCGACATGAACACA 1575
Db 241 GAGGGGTGTAGGCAATAAAAGTTATTTTATTTTTATTTTTATCTAGTGAATATA 300

Qy 1576 ATTTCTCTCACCTGGCTCCCAAAATTTAAGATGTGACACAGGCCTGTGGGTATCTCAG 1635
Db 301 ATTTTTTTTTTATTTTGGTTTTTAAATTTAAAGATGTGGAATTAAGGTTTGTGGGTATT 360

Qy 1636 GGGCAAGGAGAGCCCTGGGGTCAGTGACACTGTGAGGCCAACCATGCACATCCACAAAGG 1695
Db 361 GGGTAAGGAGAGTTTGGGGTTAGTGATTTTGGTGGTTAATTAATGATTTTATAAAGG 420

Qy 1696 GAGCATTTGGAAATGAAGGACTAGCTCTCATGTATCAGGTTTAAGACAAGGGAGAGCTGG 1755
Db 421 GAGTATTTGGAATGAAGGATTAGTTTTTATGTTATAGTTTAAGGTAAGGAGAGTTGG 480

Qy 1756 CCAGGGAACAGCATTTTGCACAGAGAGGGGAATGTAGCAACAGCAGGGCCCTCTAGGCC 1815
Db 481 TTAGGGATAGTAGTTGTATAGTAGAGGGGAATGTAGTAATAGTAGGTTTTTTTAGGTTT 540

Qy 1816 CATCTTCCATTTCTTAGTAAAGAGCANTTCTCAGACTCCACAGCGGAGGACTCAGC 1875
Db 541 TATTTTTTTATTTTGTAGTAAGAAGATATTTTTTTAGATTTTTTGGCGGAGGATGAGT 600

Qy 1876 CTAGCCTTCAGCAACCAAGGTTCTCTGGGACCCAAAAGTTTATGGGAGAAGGGCAAGAC 1935
Db 601 TTAGTTTTTAGTAATTAAGTTTTTTTTTGGGATTTTAAAGTTTATGGGAGAAGGTTAAGAT 660

Qy 1936 TTCAATGGGAAGAGAAAGGAGCCCTGGGTAGAAACCGCTTGGTGTGTTTCTCTTTGGCC 1995
Db 661 TTTATGGGAAGAGAAAGGAGGTTTTTGGGTAGAAACGTTTGGTGTGTTTTTTTTTGGTT 720

Qy 1996 TTTAAGACAAGCGCTCATCTTGGCTCTACCTCCTCATAGGCTTAGGGTTTGGCACAAC 2055
Db 721 TTTAAGATAAAGCGTTTTATTTTGTTTTTTATTTTTTGTATAGGTTTGAAGGTTTTGTTAAT 780

Qy 2056 ACATGTGGCTACAGGTGGAGGGAAGAGGACTCCTCTCCACAGTGCCTATGTTACAGAA 2115
Db 781 ATATTGTGGTTATAGGTGGAGGGAAGAGGATTTTTTTTTTTAGAGTGTATGTTTAGGAA 840

Qy 2116 GTTTCTTTTAAACCCATATGGCCCCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGAACA 2175
Db 841 GTTTTTTTTAAATTTTATATGTTTAAAGAGTAGTTCGTAGGAGGTTTTTTTAAAGACGGAATA 900

Qy 2176 AGTAATTTTACAGTTCTACTGGGGTTTCTGCCCAACCGTCCCAAGGTGGGCGAGGCTAGG 2235
Db 901 AGTAATTTTATAGTTTATTTAGGGGTTTTTGTGTTATCGTTTTTAAAGTGGGAGGTTTAGG 960

```

RESULT 4  
US-10-257-166-151  
; Sequence 151, Application US/10257166











482	QY	CTTCTATGAACAAGATGCTGCGACCAACCTCAACATTAAAGATCAATTACCGGCTGTGAC	541
518	Db	CTTCTATGAACAAGATGCTGCGACCAACCTCAACATTAAAGATCAATTACCGGCTGTGAC	577
542	QY	AGAGCTGCTGTGTCAGGCCCATAGGAAGGAGCTGGAGGCGCTCCGATGGTGCCAGCA	601
578	Db	AGAGCTGCTGTGTCAGGCCCATAGGAAGGAGCTGGAGGCGCTCCGATGGTGCCAGCA	637
602	QY	TCAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGGCAACCCGAGGGGCCACGCAACTC	661
638	Db	TGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGGCAACCCGAGGGGCCACGCAACTC	697
662	QY	TTTCGAAAACCTGCTGGTGCTGAGTCTGTGTGGGGCACCCCAACAGCACACCCCTCTTCCC	721
698	Db	TTTCGAAAACCTGCTGGTGCTGAGTCTGTGTGGGGCACCCCAACAGCACACCCCTCTTCCC	757
722	QY	TCAGAGGCGCGTGGGCGACAGAGGGAGCGGGGCTTTGGCCCTGCTGCTGTCTCTCTG	781
758	Db	TCAGAGGCGCGCTGGGCGACAGAGGGAGCGGGGCTTTGGCCCTGCTGCTGTCTCTCATG	817
782	QY	TGATGACCCCTATTGAGTATCAGTAGTACCACTACTCCCTCGCTGGCCCTGAGAGCGGCTC	841
818	Db	TGATGACCCCTATTGAGTATCAGTAGTACCACTACTCCCTCGCTGGCCCTGAGAGCGGCTC	877
842	QY	TGCTGTCTATCTCAAGCAGCCCTGCTGCCACGCGCTCCACCTCGAGTGGTCTTCTTCAG	901
878	Db	TGCTGTCTATCTCAAGCAGCCCTGCTGCCAGCCCTCCACCTGGAGTGGTCTTCTTCAG	937
902	QY	CCTGTTTCCCCCAGCCAGGCGCTCTACGACCCCCACGATGTGCCGCAAGCACTGTCTCA	961
938	Db	CCTGTTTCCCCCAGCCAGGCGCTCTACGACCCCCACGATGTGCCGCAAGCACTGTCTCA	997
962	QY	CCATCCCGGACCCACAGACAACAGCAGGGGCTGGAGTCCAGGCCA	1007
998	Db	CCATCCCGGACCCACAGACAACAGCAGGGGCTGGAGTCCAGGCCA	1043

RESULT 8  
US-10-363-616-74  
; Sequence 74, Application US/10363616  
; Publication No. US20040044181A1  
; GENERAL INFORMATION:

Query Match 30.7%: Score 1000.2: DB 12: Length 1054:

```

RESULT 9
US-09-817-198A-28
; Sequence 28, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-817-198A-28

```

## RESULT 9

Query Match 18.4%; Score 599; DB 9; Length 601;  
 Best Local Similarity 99.7%; Pred. No. 1e-154;  
 Matches 599; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1076 TTCTGTGACCCCTCCCTCCGGTCGGTTCGTATCAAAGTCTCTCAAACCCCTCCCCG 1135  
 1 TTCTGTGACCCCTCCCTCCGGTCGGTTCGTATCAAAGTCTCTCAAACCCCTCCCCG 60

Qy	1076	TTCTCTGACCCCTCCCTCCGGTGCCTTTTCGTATCAAAAGCTCTCAAACCCCTGCCCG	1135
Db	1	TTCTCTGACCCCTCCCTCCGGTGCCTTTTCGTATCAAAAGCTCTCAAACCCCTGCCCG	60
Qy	1136	TGTGTCTGTGTGTGCAGCTCGCTCTTTCCTTCTTCTTAAGCTATCCAAGGGGATGGA	1195
Db	61	TGTGTCTGTGTGTGCAGCTCGCTCTTTCCTTCTTCTTAAGCTATCCAAGGGGATGGA	120
Qy	1196	CCAGGCTGTGGGAGGTTCACCCCTTGATCCAGGAAGAACCTTCACCCCTGCCTCGT	1255
Db	121	CCAGGCTGTGGGAGGTTCACCCCTTGATCCAGGAAGAACCTTCACCCCTGCCTCGT	180





QY 643 CCGAGGCCAGCGAACTCTTGAACACCTGCTGCTGAGTCTCTGCTGAGTGGGACACCC 702  
Db 146 CCGAGGCCAGCGAACTCTTGAACACCTGCTGCTGAGTCTCTGCTGAGTGGGACACCC 205  
QY 703 ACAGCAGACCCCTCTCCCTCAGAGGCCGCTGGGACAGAGGGAGCCGGGGCTTTGCC 762  
Db 206 ACAGCAGACCCCTCTCCCTCAGAGGCCGCTGGGACAGAGGGAGCCGGGGCTTTGCC 265  
QY 763 CTGCTGCTGCTCTCTGCTGATGACCTATTGAGTATCAGTAGCCACTACTCCCTGTC 822  
Db 266 CTGCTGCTGCTCTCTGCTGATGACCTATTGAGTATCAGTAGCCACTACTCCCTGTC 325  
QY 823 CTGSCCTGAGAGGGCTCTGCTGATCTCAAGCAGCCCTGTCGCCAGCCGCTGCACC 882  
Db 326 CTGSCCTGAGAGGGCTCTGCTGATCTCAAGCAGCCCTGTCGCCAGCCGCTGCACC 385  
QY 883 CTGAGTGGTCTTCTTCAGCCTGTTTCCCGACACAGGGCTGTCAGACCCCGACGATG 942  
Db 386 CTGAGTGGTCTTCTTCAGCCTGTTTCCCGACACAGGGCTGTCAGACCCCGACGATG 445  
QY 943 TGGCGAAGCACTGCTCAGCATCCCGACCCAGCAGACACAGCCAGGGCTGGAGTCCA 1002  
Db 446 TGGCGAAGCACTGCTCAGCATCCCGACCCAGCAGACACAGCCAGGGCTGGAGTCCA 505  
QY 1003 GGGCCTCTTCTCTCTCTGCTGACCCCTGCC 1091  
Db 566 CCCCCACTCTCTCTCTGACCCCTGCC 594

## RESULT 12

US-09-817-198A-30  
; Sequence 30, Application US/09817198A  
; Patent No. US20020146758A1  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001188  
; CURRENT APPLICATION NUMBER: US/09/817.198A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-817-198A-30

Query Match 16.6%; Score 540; DB 9; Length 601;  
Best Local Similarity 99.6%; Pred. No. 2e-138;  
Matches 540; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2682 GCCAGGACCTTGAGAGTTATTGGGAGTGGCTCCAACTGTGGACCACTATTTCAGC 2741  
Db 1 GCCAGGACCTTGAGAGTTATTGGGAGTGGCTCCAACTGTGGACCACTATTTCAGC 60  
QY 2742 TTTCCCTGAAGATCAGGAGGGTGCATTTCATTTCTCTCTCTAGCCCCCTCAGAA 2801  
Db 61 TTTCCCTGAAGATCAGGAGGGTGCATTTCATTTCTCTCTAGCCCCCTCAGAA 120  
QY 2802 AGAAGGACTATATTGTACTGTACCTAGGGGTTCTGGAAGGGAAAAATGGAATCAGGA 2861  
Db 121 AGAAGGACTATATTGTACTGTACCTAGGGGTTCTGGAAGGGAAAAATGGAATCAGGA 180  
QY 2862 TTCTATAGACTGATAGGCCCTTATCCAGAGGCCATGATCGGAAAAAGGTATGGAGCAG 2921  
Db 181 TTCTATAGACTGATAGGCCCTTATCCAGAGGCCATGATCGGAAAAAGGTATGGAGCAG 240

QY 2922 AAGGAGAAATGGGATTTTAGGGTGACGCTACGCTCACCTAAACTTTTGGTGGCCTGGGG 2981  
Db 241 AAGGAGAAATGGGATTTTAGGGTGACGCTACGCTCACCTAAACTTTTGGTGGCCTGGGG 300  
QY 2982 CATGCTTTGAGGCCAGAGACTGTTAAGCAGGCTCTGCTGGCTGTTTACTGCTCACCACCT 3041  
Db 301 YATGCTTTGAGGCCAGAGACTGTTAAGCAGGCTCTGCTGGCTGTTTACTGCTCACCACCT 360  
QY 3042 CTGACCTGCTGCTGAGACTCCATCCAGCCCGCAGCGCCACCTGCTCCCTGAGCCTC 3101  
Db 361 CTGACCTGCTGCTGAGACTCCATCCAGCCCGCAGCGCCACCTGCTCCCTGAGCCTC 420  
QY 3102 CACTATCTCCTGTGAGCGGTGAACCTTCGTGTACTGTCTCGGGTCCATATATGAATTG 3161  
Db 421 CACTATCTCCTGTGAGCGGTGAACCTTCGTGTACTGTCTCGGGTCCATATATGAATTG 480  
QY 3162 TGAGCAGGGTTCATCTATTTTAAACACAGATGTTTACAAATAAAGATTATTTCAAACCA 3221  
Db 481 TGAGCAGGGTTCATCTATTTTAAACACAGATGTTTACAAATAAAGATTATTTCAAACCA 540  
QY 3222 CC 3223  
Db 541 CC 542

## RESULT 13

US-09-764-868-507  
; Sequence 507, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PTZ32  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 507  
; LENGTH: 566  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: (484)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (493)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (538)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (563)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-868-507

Query Match 16.4%; Score 534.2; DB 9; Length 566;  
Best Local Similarity 98.9%; Pred. No. 7.9e-137;  
Matches 533; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 2 GCCCGCTGCCCGCCCGCAGTTCCCGGCCCGCTGGCCCGCAGTCATGCGGAGCAGTACGA 61  
Db 28 GCCCGCTGCCCGCCCGCAGTTCCCGGCCCGCTGGCCCGCAGTCATGCGGAGCAGTACGA 87  
QY 62 TGTGCTGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTGCTGTG 121  
Db 88 TGTGCTGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTGCTGTG 147  
QY 122 CCGGTTCCAGCAACAGATTCCACTCTCGCACATCTCCACCATCGTGTGACTTTAA 181  
Db 148 CCGGTTCCAGCAACAGATTCCACTCTCGCACATCTCCACCATCGTGTGACTTTAA 207



Qy 182 GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACACTGCAGG 241  
Db 208 GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTGGGACACTGCAGG 267  
Qy 242 GCAGGAGAGATACCAAGACCATCAAAAGTGGGATACAGATCTGGGACACTGCAGG 301  
Db 268 GCAGGAGAGATACCAAGACCATCAAAAGTGGGATACAGATCTGGGACACTGCAGG 327  
Qy 302 GGTCTATCACATTAGCAGGAGCGCTCTTACAGCACATCATGAAGTGGGTCAAGTACGT 361  
Db 328 GGTCTATGACATTAGCAGGAGCGCTCTTACAGCACATCATGAAGTGGGTCAAGTACGT 387  
Qy 362 GGATGAGTACGACACAGAGGCGCTCCAGAAATCTTATTGGGAATAGGCTGATGAGGA 421  
Db 388 GGATGAGTACGACACAGAGGCGCTCCAGAAATCTTATTGGGAATAGGCTGATGAGGA 447  
Qy 422 GCAGAAACGGCAGGTGGGAAAGAGAGCAGGAGCTGGGAGGAGTATGGCATGGA 481  
Db 448 GCAGAAACGGCAGGTGGGAAAGAGAGCAGGAGCTGGGAGGAGTATGGCATGGA 507  
Qy 482 CTTCTATGAACAAGTGGTCCACCACTCAAACTCAAAAGTCAATTCAGCGCTCTGA 540  
Db 508 CTTCTATGAACAAGTGGTCCACCACTCAAACTCAAAAGTCAATTCAGCGCTCTGA 566

## RESULT 14

US-09-817-198A-31  
; Sequence 31, Application US/09817198A  
; Patent No. US20020146758A1  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001188  
; CURRENT APPLICATION NUMBER: US/09/817.198A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 31  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-817-198A-31

Query Match 15.9%; Score 516.6; DB 9; Length 601;  
Best Local Similarity 99.8%; Pred. No. 6e-132;  
Matches 516; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 2707 GGCAGTGGCTCCAATCTGTGACACAGTATTTTCAGCTTTCCCTGGAAGATCAGGCGGGTGC 2766  
Db 1 GGCAGTGGCTCCAATCTGTGACACAGTATTTTCAGCTTTCCCTGGAAGATCAGGCGGGTGC 60  
Qy 2767 CATTCATTGCTTTTCTCTCTAGCCCTCAGGAAAGAGACTATATTTGCTACTGTACC 2826  
Db 61 CATTCATTGCTTTTCTCTCTAGCCCTCAGGAAAGAGACTATATTTGCTACTGTACC 120  
Qy 2827 CTAGGGGTCTGGAAGGAAACATGGAATCAGATCTATAGACTATAGCCCTATCC 2886  
Db 121 CTAGGGGTCTGGAAGGAAACATGGAATCAGATCTATAGACTATAGCCCTATCC 180  
Qy 2887 ACAAGGGCCATGACTGGGAAAGGTATGGGAGCAGAGGAGAAATTTGGATTTAGGGTGC 2946  
Db 181 ACAAGGGCCATGACTGGGAAAGGTATGGGAGCAGAGGAGAAATTTAGGGTGC 240  
Qy 2947 AGCTACGCTACCCCTAAACTTTTGGTGGCTGGGCGATGCTTCGAGGCCACAGACTGTAA 3006  
Db 241 AGCTACGCTACCCCTAAACTTTTGGTGGCTGGGCGATGCTTCGAGGCCACAGACTGTAA 300  
Qy 3007 GCAGGCTCTGCTGGCTGTGTTACTGCTGCACCACTCTGCACTGCTGTGAGACTCCA 3066  
Db 301 SCAGGCTCTGCTGGCTGTGTTACTGCTGCACCACTCTGCACTGCTGTGAGACTCCA 360

Qy 3067 TCCAGCCCCAGGACGCGACCTGCTCTGAGCCTCCACTATCTCTCTGACGGGTGAAC 3126  
Db 361 TCCAGCCCCAGGACGCGACCTGCTCTGAGCCTCCACTATCTCTCTGACGGGTGAAC 420  
Qy 3127 TTCGTGACTGTGTCTCGGTCCATATATGAATTTGTGAGCAGGGTCAATCTATTTAAAC 3186  
Db 421 TTCGTGACTGTGTCTCGGTCCATATATGAATTTGTGAGCAGGGTTCATCTATTTTAAAC 480  
Qy 3187 ACAGATGTTTACAAAATAAAGATTATTTTCAAAACCACC 3223  
Db 481 ACAGATGTTTACAAAATAAAGATTATTTTCAAAACCACC 517

## RESULT 15

US-09-920-300A-303  
; Sequence 303, Application US/09920300A  
; Patent No. US20020136728A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547  
; CURRENT APPLICATION NUMBER: US/09/920.300A  
; CURRENT FILING DATE: 2001-07-31  
; NUMBER OF SEQ ID NOS: 1789  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 303  
; LENGTH: 481  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 417, 461  
; OTHER INFORMATION: n = A,T,C or G  
US-09-920-300A-303

Query Match 14.7%; Score 477.4; DB 9; Length 481;  
Best Local Similarity 99.4%; Pred. No. 3.6e-121;  
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 2582 GTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCACAACCTGGG 2641  
Db 1 GTTGCTTCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCACAACCTGGG 60  
Qy 2642 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGACTTGAGAAGTTA 2701  
Db 61 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGACTTGAGAAGTTA 120  
Qy 2702 TATTGGCAGTGGTCCAAATCTGTGACCAAGTATTTCCAGCTTTCCCTGAAAGATCAGGAG 2761  
Db 121 TATTGGCAGTGGTCCAAATCTGTGACCAAGTATTTCCAGCTTTCCCTGAAAGATCAGGAG 180  
Qy 2762 GGTGCCAATTCATTGCTCTTCTCTAGCCCTCAGGAAAGAGGACTATATTTGTTACT 2821  
Db 181 GGTGCCAATTCATTGCTCTTCTCTAGCCCTCAGGAAAGAGGACTATATTTGTTACT 240  
Qy 2822 GTACCTTAGGGGTTCTGGAAGGGGAAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 2881  
Db 241 GTACCTTAGGGGTTCTGGAAGGGGAAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 300  
Qy 2882 TATCCACAGGGCCATGACTGGGAAAGGTATGGGACAGAGAGAAATTTGGGATTTTAG 2941  
Db 301 TATCCACAGGGCCATGACTGGGAAAGGTATGGGACAGAGAGAAATTTGGGATTTTAG 360  
Qy 2942 GGTGCAGCTACGCTCACCCCTAAACTTTTGGTGGCTGGGCGATGCTTTGAGGCCACAGCT 3001  
Db 361 GGTGCAGCTACGCTCACCCCTAAACTTTTGGTGGCTGGGCGATGCTTTGAGGCCACAGCT 420  
Qy 3002 GTTAAAGCAGGCTCTGCTGGGCTGTTTACTGCTCACCACCTCTGCACCTGCTGCTTGGAGA 3061

Db 421 GTTAACCAAGGCTCTGCTGGGCTGTTTACTGTCACCACTNTGCACCTGCTGCTTGAGA 480

Qy 3062 C 3062

Db 481 C 481

Search completed: March 21, 2004, 05:17:48  
Job time : 1091 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 04:56:29 ; Search time 7314 Seconds  
(without alignments)  
13297.938 Million cell updates/sec

Title: US-09-817-198C-1

Perfect score: 3257

Sequence: 1 tgcgcgtgccgcgcgcag.....aaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2874.2	88.2	3151	11	CNSLTI18T
2	2690.4	82.6	2700	11	BC014511
3	947.8	29.1	1081	12	BM551913
4	883.4	27.1	956	13	BM551913

5	856	26.3	957	13	BQ937298
6	847.2	26.0	915	13	BX397020
7	840.2	25.8	908	13	BQ167774
8	837.8	25.7	918	13	BX326279
9	810.6	24.9	901	13	BQ220195
10	807.8	24.8	929	13	BX348287
11	801.6	24.6	924	12	B1118034
12	799.2	24.5	1046	13	BQ277739
13	797.2	24.5	840	13	BQ557215
14	792.2	24.3	820	13	BQ849866
15	792.2	24.3	863	13	BQ557274
16	781.2	24.0	1027	12	BM476862
17	779.6	23.9	907	13	BQ541549
18	769.4	23.6	922	13	BQ844769
19	767.4	23.6	878	12	BQ676934
20	758.4	23.3	928	13	BQ151020
21	755.4	23.2	767	12	BM977746
22	746.2	22.9	763	12	BM975016
23	746	22.9	990	10	BE799964
24	743.6	22.8	793	14	CA411919
25	742	22.8	756	13	BX093757
26	733.8	22.5	738	13	BQ172073
27	723.6	22.2	733	13	BX331206
28	722	22.2	732	12	BM045331
29	721.8	22.2	749	12	BQ769088
30	716.8	22.0	766	12	BM049329
31	704.6	21.6	729	13	BX352668
32	704.6	21.6	907	13	BX388608
33	698.2	21.4	1116	13	BQ067682
34	697	21.4	744	12	BQ697415
35	695.2	21.3	703	13	BQ687540
36	689.8	21.2	845	12	BQ821168
37	689.4	21.2	783	12	BQ760915
38	687.2	21.1	887	12	BQ766350
39	687.2	21.1	935	13	BQ876547
40	664	20.4	825	14	CD643455
41	662	20.3	899	13	BQ552993
42	644.2	19.8	891	10	BE883791
43	638	19.6	639	13	BQ188717
44	632.2	19.4	637	14	CA424721
45	618.2	19.0	647	10	BF966846

ALIGNMENTS

RESULT 1	CNSLTI18T	3151 bp	mRNA	linear	HTC 18-JUN-2003
LOCUS	human full-length cDNA clone CS0D1028YM11 of Placenta of Homo sapiens (human).				
DEFINITION	human full-length cDNA clone CS0D1028YM11 of Placenta of Homo sapiens (human).				
ACCESSION	BX248046				
VERSION	BX248046.1	GI:28193221			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 3151)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (30-JUN-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned				

into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

source  
Location/Qualifiers  
1..3151  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS01028YM11"  
/tissue="Placenta"  
/notes="Cot 25-normalized-vector pCMVSPORT\_6"  
<1..504  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD62353.1"  
/db\_xref="GI:28193222"  
/translation="GVDFPKMTIEVDGIKVIQIWDTAGQERYQTITKQYVRAQGLF  
LVYDISERSYQHIMKWSVDVEVDGATSLPGCGEGASPGKARRGPDGKANARKLCL  
POPWMKTSYGHQKASRRSLGILRMSRNGRWEESKGSWRRSMAMWTSKQVPAFTST  
LKSHSRV"

CDS

ORIGIN

Query Match 88.2%; Score 2874.2; DB 11; Length 3151;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 3017; Conservative 0; Mismatches 3; Indels 131; Gaps 1;  
Qy 168 GGTGTTGACATTTAAGATGAAGACATAGAGGTAGACGCGCATCAAAGTGGCGATACAGATC 227  
Db 1 GGTGTTGACATTTAAGATGAAGACATAGAGGTAGACGCGCATCAAAGTGGCGATACAGATC 60  
Qy 228 TGGGACACTGCGAGCGGAGAGATACAGACCATCAAAAGCAGTACTATCGCGGGGCC 287  
Db 61 TGGGACACTGCGAGCGGAGAGATACAGACCATCAAAAGCAGTACTATCGCGGGGCC 120  
Qy 288 CAGGGGATATTTTGGTCTATGACATTTAGCAGCGAGCGCTTTACAGGACATCATGAAG 347  
Db 121 CAGGGGATATTTTGGTCTATGACATTTAGCAGCGAGCGCTTTACAGGACATCATGAAG 180  
Qy 348 TGGGTGAGTACGTGGATGA----- 367  
Db 181 TGGGTGAGTACGTGGATGAGTAGGATAGCCACCTCACTGCCGGGGTGTGGAGAGGT 240  
Qy 368 ----- 367  
Db 241 GCCTCACGGGGAAGGCAAGCGAGGCGCAGATGGGAAGGCAAAATGCTTCCAGGAAGCTT 300  
Qy 368 -----GTAGGCACCAAGAGGCGTCCAGAGATCC 396  
Db 301 TGCCTTCCACAGCCCTGGATGAAGACCTCTGGTACGCACCAAGAGCGCTCCAGAGATCC 360  
Qy 397 TTATTGGGAATAAGGCTGATGAGGAGCAGAAACGCGAGGTGGGAAGAGAGCAAGGCGAGC 456  
Db 361 TTATTGGGAATAAGGCTGATGAGGAGCAGAAACGCGAGGTGGGAAGAGAGCAAGGCGAGC 420  
Qy 457 AGCTGGCGAAGGAGTATGGCATGGAATTTATGAAACAAGTGCCTGCACCAACCTCAACA 516  
Db 421 AGCTGGCGAAGGAGTATGGCATGGAATTTATGAAACAAGTGCCTGCACCAACCTCAACA 480  
Qy 517 TTAAGAGTCAATTCAGCGTCTGACAGCTGTGTGCTGCAAGGCCATATAGGAAGAGCTGG 576  
Db 481 TTAAGAGTCAATTCAGCGTCTGACAGAGCTGTGTGCTGCAAGGCCATATAGGAAGAGCTGG 540  
Qy 577 AAGSCCTCCGATCGGTCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGG 636  
Db 541 AAGSCCTCCGATCGGTCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGG 600  
Qy 637 GCAACCCGAGGGCCCGCAACTCTTCGAAACCTGCTGTGTGCTGAGTCCCTGTGTGGGG 696  
Db 601 GCAACCCGAGGGCCCGCAACTCTTCGAAACCTGCTGTGTGCTGAGTCCCTGTGTGGGG 660  
Qy 697 CACCCCAACAGACACCCCTCTTCCTCAGAGAGGCCCGTGGGCAGACAGGGAGCCGGGGC 756  
Db 661 CACCCCAACAGACACCCCTCTTCCTCAGAGAGGCCCGTGGGCAGACAGGGAGCCGGGGC 720

Qy 757 TTGCCCCCTGCTGCTCTCTCGTGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCC 816  
Db 721 TTGCCCCCTGCTGCTCTCTCGTGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCC 780  
Qy 817 CCTGCTGCGCCCTGAGAGCGGCTCTGCTGTATCTCAAGAGAGCCCTGTGCCAGGCCG 876  
Db 781 CCTGCTGCGCCCTGAGAGCGGCTCTGCTGTATCTCAAGAGAGCCCTGTGCCAGGCCG 840  
Qy 877 TCCACCCCTGGAGTGTCTCTTCTCAGCGTGTTCCTCCAGGCCACAGCGGCTGTACGACCC 936  
Db 841 TCCACCCCTGGAGTGTCTCTTCTCAGCGTGTTCCTCCAGGCCACAGCGGCTGTACGACCC 900  
Qy 937 ACGATGTGCCGCAAGCACTGTCTCACCATCCCGCACCCACAGACAAACAGCGAGGCTGG 996  
Db 901 ACGATGTGCCGCAAGCACTGTCTCACCATCCCGCACCCACAGACAAACAGCGAGGCTGG 960  
Qy 997 AGTCCAGGCGCACTTTTCAAGTGTCTCTTCTCGGTGCAATCGTGTCTCTTCTGTCTTTTC 1056  
Db 961 AGTCCAGGCGCACTTTTCAAGTGTCTCTTCTCGGTGCAATCGTGTCTCTTCTGTCTTTTC 1020  
Qy 1057 TCTCTTCCCCCACTTCTCTTCTCTGACCCCTCCCTCCGGTGGTTCGTATCAAGACT 1116  
Db 1021 TCTCTTCCCCCACTTCTCTTCTCTGACCCCTCCCTCCGGTGGTTCGTATCAAGACT 1080  
Qy 1117 CCTCAAAACCCCGTCCCGGTGTCTCTGTGTGTGACAGCTCGCTCTTCTTCTTCTTCTTA 1176  
Db 1081 CCTCAAAACCCCGTCCCGGTGTCTCTGTGTGTGACAGCTCGCTCTTCTTCTTCTTCTTA 1140  
Qy 1177 AGTATCAAAAGGGATGAGCCAGGCTCTGTTGGGAGGTTCCACCTTTGGATCCAGGAAGA 1236  
Db 1141 AGTATCAAAAGGGATGAGCCAGGCTCTGTTGGGAGGTTCCACCTTTGGATCCAGGAAGA 1200  
Qy 1237 ACCTCCACCCCTGCTGTTGGGTGGGCAAGGCTACAGGGTGTCTTCTTCTTCTTCCCC 1296  
Db 1201 ACCTCCACCCCTGCTGTTGGGTGGGCAAGGCTACAGGGTGTCTTCTTCTTCTTCCCC 1260  
Qy 1297 ACCCCCACTGTCTCTATGTGCCATGGGCTGTCTCCCAAGTGAACCTGCGAAAGTGGAGC 1356  
Db 1261 ACCCCCACTGTCTCTATGTGCCATGGGCTGTCTCCCAAGTGAACCTGCGAAAGTGGAGC 1320  
Qy 1357 ATCGAGGTAGAGGGGAAAACAGCAACCGGGGAGTCTCTGAGCGCTGGGCTGCGCTTACCT 1416  
Db 1321 ATCGAGGTAGAGGGGAAAACAGCAACCGGGGAGTCTCTGAGCGCTGGGCTGCGCTTACCT 1380  
Qy 1417 ACCATTTCCCGACAGAGCTTTGCTTGGCTGCGCGCTGCTCTTTGGGGAAC 1476  
Db 1381 ACCATTTCCCGACAGAGCTTTGCTTGGCTGCGCGCTGCTCTTTGGGGAAC 1440  
Qy 1477 TGAGCTCAGAGGCGAGGTGCTTTCCAGAGAAAGAAACAAAATGAGGGGTGGCAGGATAAAAA 1536  
Db 1441 TGAGCTCAGAGGCGAGGTGCTTTCCAGAGAAAGAAACAAAATGAGGGGTGGCAGGATAAAAA 1500  
Qy 1537 GTCACTCTCAATCTCTACCTCCCATGAGCATGAACAATAATTTCTCTCACTGGCTCCC 1596  
Db 1501 GTCACTCTCAATCTCTACCTCCCATGAGCATGAACAATAATTTCTCTCACTGGCTCCC 1560  
Qy 1597 AAATTTAAAGATGTGGACCAAGGCTGTGGTACTCCAGGGGCAAGAGAGCCCTGGGGT 1656  
Db 1561 AAATTTAAAGATGTGGACCAAGGCTGTGGTACTCCAGGGGCAAGAGAGCCCTGGGGT 1620  
Qy 1657 CAGTGACACTGTTCAGGCGCAACCTGCACTCAAAAGGGAGCATTTTGGAAATGAAGGAC 1716  
Db 1621 CAGTGACACTGTTCAGGCGCAACCTGCACTCAAAAGGGAGCATTTTGGAAATGAAGGAC 1680  
Qy 1717 TAGCTCTTATGTATCAGGTTTAAGAGCAAGGAGAGTGGCCAGGACAGAGTTTGACA 1776  
Db 1681 TAGCTCTTATGTATCAGGTTTAAGAGCAAGGAGAGTGGCCAGGACAGAGTTTGACA 1740  
Qy 1777 GCAGAGGGGAAATGTAGCAACAGCAGGGCTCTTAGGCCCATCTTCCATTTCTTAGGTAA 1836  
Db 1741 GCAGAGGGGAAATGTAGCAACAGCAGGGCTCTTAGGCCCATCTTCCATTTCTTAGGTAA 1800

Qy 1837 GAAGAGCAATTCCTCAGACTCCAGCGGAGGACTGAGCCCTAGCCTTCCAGCAACCAAGGT 1896  
Db 1801 GAAGAGCAATTCCTCAGACTCCAGCGGAGGACTGAGCCCTAGCCTTCCAGCAACCAAGGT 1860  
Qy 1897 TCTCCTGGGACCCAAAGTTTATGGAGAGGCAAGCACTTCATGGGAAGAGAGGAA 1956  
Db 1861 TCTCCTGGGACCCAAAGTTTATGGAGAGGCAAGCACTTCATGGGAAGAGAGGAA 1920  
Qy 1957 GGCCCTGGGTAGAAACGCTTGGTGTCTTCTTCTTGGCTTTAAGACAAAGCGCTCATCT 2016  
Db 1921 GGCCCTGGGTAGAAACGCTTGGTGTCTTCTTCTTGGCTTTAAGACAAAGCGCTCATCT 1980  
Qy 2017 TGCCCTCTACCTCTGTAGAGCTTGAAGTTTGCACCAACCACTGTGGCTACAGGTGAG 2076  
Db 1981 TGCCCTCTACCTCTGTAGAGCTTGAAGTTTGCACCAACCACTGTGGCTACAGGTGAG 2040  
Qy 2077 GGAAGAGGACTCCTTCTCCAGAGTGCTATGTTTCCAGGAGTTTCTTAAACCCATATGGC 2136  
Db 2041 GGAAGAGGACTCCTTCTCCAGAGTGCTATGTTTCCAGGAGTTTCTTAAACCCATATGGC 2100  
Qy 2137 CCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGAACCAAGTAATTTACCAAGTTCTACTG 2196  
Db 2101 CCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGAACCAAGTAATTTACCAAGTTCTACTG 2160  
Qy 2197 GGGTTCTGCCCCACCGTCCCAAGGTGGCGAGGCTAGGAAGAGGTCATTTTAAAGCCA 2256  
Db 2161 GGGTTCTGCCCCACCGTCCCAAGGTGGCGAGGCTAGGAAGAGGTCATTTTAAAGCCA 2220  
Qy 2257 CACATTAGCTCAGCTGGTGGCTGCAGCCAAACCAAGAACTGGGTGTGAGTATTCATC 2316  
Db 2221 CACATTAGCTCAGCTGGTGGCTGCAGCCAAACCAAGAACTGGGTGTGAGTATTCATC 2280  
Qy 2317 AACTAAGAACCAAAATCAGGCGCTCATATGTGAAGATAAGAACCTCACTTCTTACT 2376  
Db 2281 AACTAAGAACCAAAATCAGGCGCTCATATGTGAAGATAAGAACCTCACTTCTTACT 2340  
Qy 2377 CCTCCAAAGAAAGTGGGGAAGAACCAATCAAACTTTCTTCCCTGACTTACCAACCAAGG 2436  
Db 2341 CCTCCAAAGAAAGTGGGGAAGAACCAATCAAACTTTCTTCCCTGACTTACCAACCAAGG 2400  
Qy 2437 AAAACAGCAGAGAGGGTGGCTCAGGACTTAGGACAGGGTATAGCTTAGTGTGGAA 2496  
Db 2401 AAAACAGCAGAGAGGGTGGCTCAGGACTTAGGACAGGGTATAGCTTAGTGTGGAA 2460  
Qy 2497 GCAAGGAGAGCAGCAAGTGTGAATCACTGCTTAATCAGAAAAAGGAGACAGTAACTCT 2556  
Db 2461 GCAAGGAGAGCAGCAAGTGTGAATCACTGCTTAATCAGAAAAAGGAGACAGTAACTCT 2520  
Qy 2557 AGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTTGAAGATGGGACTCCTTGGGTATCA 2616  
Db 2521 AGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTTGAAGATGGGACTCCTTGGGTATCA 2580  
Qy 2617 AGACTATGCCACATCAGCTGGGGCTAGGAAGTAGGTATGCCAGCCCTCAAGTCTGT 2676  
Db 2581 AGACTATGCCACATCAGCTGGGGCTAGGAAGTAGGTATGCCAGCCCTCAAGTCTGT 2640  
Qy 2677 CTTCCAGCCAGGAGCTTGAGAGAGTTATATTGGGAGTGGCTCCAACTCTGTGGACCAAGTAT 2736  
Db 2641 CTTCCAGCCAGGAGCTTGAGAGAGTTATATTGGGAGTGGCTCCAACTCTGTGGACCAAGTAT 2700  
Qy 2737 TCAGCTTTCCCTGAAGATCAGGCGAGGTCGCAATTCATTTCTTCTCTAGCCCCCTC 2796  
Db 2701 TCAGCTTTCCCTGAAGATCAGGCGAGGTCGCAATTCATTTCTTCTCTAGCCCCCTC 2760  
Qy 2797 AGGAAAGAGCACTATATTGTACTGTACCTTAGGGTTCTGGAAGGGAACATGGAA 2856  
Db 2761 AGGAAAGAGCACTATATTGTACTGTACCTTAGGGTTCTGGAAGGGAACATGGAA 2820  
Qy 2857 CAGGATTTATAGACTGATGAGCCCTATCCACAGGGCCATGACTGGGAAAGTATGGG 2916  
Db 2821 CAGGATTTATAGACTGATGAGCCCTATCCACAGGGCCATGACTGGGAAAGTATGGG 2880  
Qy 2917 AGCAGAGGAGAAATGGGATTTTAGGGTGCAGTACGCTCACCCCTAAACTTTTGTGGGCC 2976

Db 2881 AGCAGAGGAGAAATGGGATTTTAGGGTGCAGTACGCTCACCCCTAAACTTTTGTGGGCC 2940  
Qy 2977 TGGGGCATGCTTTGAGGCCAGACTGTTAAGCAGGCTCTGCTGGCTGTTTACTCGTCAC 3036  
Db 2941 TGGGGCATGCTTTGAGGCCAGACTGTTAAGCAGGCTCTGCTGGCTGTTTACTCGTCAC 3000  
Qy 3037 CACCTTCGACCTGCTGCTGTTGAGACTCCATCCAGCCCCAGGACGACCTGCTCTCTGA 3096  
Db 3001 CACCTTCGACCTGCTGCTGTTGAGACTCCATCCAGCCCCAGGACGACCTGCTCTCTGA 3060  
Qy 3097 GCCTCCACTATCTCCCTGTGACGGGTGAACCTTCTGTACTGTGTCTCGGTCCTATATG 3156  
Db 3061 GCCTCCACTATCTCCCTGTGACGGGTGAACCTTCTGTACTGTGTCTCGGTCCTATATG 3120  
Qy 3157 AATTGTGAGCAGGGTTCATCTATTTTAAACA 3187  
Db 3121 AATTGTGAGCAGGGTTCATCTATTTTAAACA 3151

RESULT 2  
BC014511  
LOCUS  
DEFINITION Homo sapiens RAB15, member RAS oncogene family, mRNA (cDNA clone IMAGE:4866926), with apparent retained intron.  
ACCESSION BC014511  
VERSION BC014511.1 GI:15680292  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
1 (bases 1 to 2700)  
Srausberg, R.L., Reingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Rodrigues, Y., Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Richardson, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
2 (bases 1 to 2700)  
Srausberg, R.  
Direct Submission  
Submitted (17-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NHI-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [gcgaps-remail.nih.gov](mailto:gcgaps-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,



```
Qy 2341 CTGATATGTAAGGATAAGAACCTCCTACTTCTCTCTCAAAAGAAAGTGGGAAAGA 2400
Db 1801 CTCATATGTGAAGGATAAGAACCTCCTACTTCTCTCAAAAGAAAGTGGGAAAGA 1860
Qy 2401 ACATCAAACTTTCTCTCTGACTTACCAACACAGGAAACAGACAGAGAGGGTGGCTCA 2460
Db 1861 ACCATCAAACTTTCTCTCTGACTTACCAACACAGGAAACAGACAGAGAGGGTGGCTCA 1920
Qy 2461 GGACTTAGGACAGGGTATAGCTTAGATGGTGAAGCAAGAGGAGAGCAGGAAGTGTGAA 2520
Db 1921 GGACTTAGGACAGGGTATAGCTTAGATGGTGAAGCAAGAGGAGAGCAGGAAGTGTGAA 1980
Qy 2521 ATCACTGGCTAATGAGAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGTAGCTAGGCTGG 2580
Db 1981 ATCACTGGCTAATGAGAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGTAGCTAGGCTGG 2040
Qy 2581 AGTTGCTTCTTGAAGATGGAGCTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640
Db 2041 AGTTGCTTCTTGAAGATGGAGCTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2100
Qy 2641 GCTAGGGAAGTAGTGATGCCAGCCCTCAAGTCTGTCTTACGCCAGGACTTGAAGAATT 2700
Db 2101 GCTAGGGAAGTAGTGATGCCAGCCCTCAAGTCTGTCTTACGCCAGGACTTGAAGAATT 2160
Qy 2701 ATATTGGGCACTGCTCAATCTCTGGACCAAGTATTTTCAGCTTTTCCCTGAAGATCAGGCA 2760
Db 2161 ATATTGGGCACTGCTCAATCTCTGGACCAAGTATTTTCAGCTTTTCCCTGAAGATCAGGCA 2220
Qy 2761 GGGTGCCATTCATTTCTCTCTCTAGCCCTCTAGGAAAGAGCAATATTTGTAC 2820
Db 2221 GGGTGCCATTCATTTCTCTCTCTAGCCCTCTAGGAAAGAGCAATATTTGTAC 2280
Qy 2821 TGTACCTTAGGGTCTTGGAGGGAACATGGATCAGGATTTCTATAGCTGTAGGCC 2880
Db 2281 TGTACCTTAGGGTCTTGGAGGGAACATGGATTCAGGATTTCTATAGCTGTAGGCC 2340
Qy 2881 CTATCCCAAGGGCATGACTGGCAAGAGGTATGGGAGCAGAAGAAATTTGGATTTTA 2940
Db 2341 CTATCCCAAGGGCATGACTGGCAAGAGGTATGGGAGCAGAAGAAATTTGGATTTTA 2400
Qy 2941 GGGTGCACTACGCTACCCCTAAATTTTGTGGCTGGGCAATGTCTTGAAGGCCAGAC 3000
Db 2401 GGGTGCACTACGCTACCCCTAAATTTTGTGGCTGGGCAATGTCTTGAAGGCCAGAC 2460
Qy 3001 TGTAAAGCAGGCTCTGTGGCTGTCTGTCGTCACCACTCTGCACTGCTGTCTTGAG 3060
Db 2461 TGTAAAGCAGGCTCTGTGGCTGTCTGTCGTCACCACTCTGCACTGCTGTCTTGAG 2520
Qy 3061 ACTCCATCCAGCCCGCAGCAGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGG 3120
Db 2521 ACTCCATCCAGCCCGCAGCAGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGG 2580
Qy 3121 GTGAACCTTGTGTACTGTGTCTCGGGTCCATATATGAATTTGTAGCAGGGTTCATCTATT 3180
Db 2581 GTGAACCTTGTGTACTGTGTCTCGGGTCCATATATGAATTTGTAGCAGGGTTCATCTATT 2640
Qy 3181 TTAACACAGATGTTTACAAATTAAGATTTATTTCAACACCCCAAAAAAAAAAAAAA 3240
Db 2641 TTAACACAGATGTTTACAAATTAAGATTTATTTCAACACCAAAAAAAAAAAAAA 2700
```

RESULT 3  
BM551913  
LOCUS  
DEFINITION AGNCOURT 6575472 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5479433  
5', mRNA sequence.  
ACCESSION BM551913  
VERSION BM551913.1 GI:18789340  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 1081)

NIH-MGC http://mgi.nhl.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LICM2000 row: k column: 18

High quality sequence stop: 662.

Location/Qualifiers

1. 1081

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5479433"

/tissue\_type="astrocytoma grade IV, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_98"

/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:

ECORI; cDNA made by oligo-dT priming. Directionally

cloned into EORI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 29.1%; Score 947.8; DB 12; Length 1081;

Best Local Similarity 95.4%; Pred. No. 8.2e-133;

Matches 1029; Conservative 0; Mismatches 44; Indels 6; Gaps 5;

Qy 1178 GCTATCAAGGGGATGACCCAGGCTCTGGGGAGGTTCCACCCCTTGATCCAGAGAA 1237

Db 1 GCTATCAAGGGGATGACCCAGGCTCTGGGGAGGTTCCACCCCTTGATCCAGAGAA 60

Qy 1238 CCCTCCACCCCTGCTCTGGTGGGCAAAAGGCTACAGGGTGCTTCTCTCTTCCCCCA 1297

Db 61 CCCTCCACCCCTGCTCTGGTGGGCAAAAGGCTACAGGGTGCTTCTCTCTTCCCCCA 120

Qy 1298 CCCCACCTGCTCTCATGTGCCATGCGCTGCTCCCAAGTGACCTGCGAAAGTGGAGCA 1357

Db 121 CCCCACCTGCTCTCATGTGCCATGCGCTGCTCCCAAGTGACCTGCGAAAGTGGAGCA 180

Qy 1358 TCGAGGTAGAGGGAACACAGCAACCGGGAGTCTCGAGCTGGGGCTGCCCTACCTCTA 1417

Db 181 TCGAGGTAGAGGGAACACAGCAACCGGGAGTCTCGAGCTGGGGCTGCCCTACCTCTA 240

Qy 1418 CCCTATCCCGACACAGAGCTTTGCCCTTGTGCTGCCCGCTGCTCTTTGGGAACT 1477

Db 241 CCCTATCCCGACACAGAGCTTTGCCCTTGTGCTGCCCGCTGCTCTTTGGGAACT 300

Qy 1478 GAGCTCAGAGGCAAGTCTTTCAGAGAGGAAACAAAATGAGGGGTGGCAGGGATAAAAAG 1537

Db 301 GAGCTCAGAGGCAAGTCTTTCAGAGAGGAAACAAAATGAGGGGTGGCAGGGATAAAAAG 360

Qy 1538 TCACCTCATTTCTTACCTCCATGACGATGAACACAAATTTCTCTCCACCTGGCTCCCA 1597

Db 361 TCACCTCATTTCTTACCTCCATGACGATGAACACAAATTTCTCTCCACCTGGCTCCCA 420

Qy 1598 AATTTAAAGATGTGACCAAGGCTGTGGGTACTTCCAGGGGCAAGAGAGCCCTGGGGTC 1657

Db 421 AATTTAAAGATGTGACCAAGGCTGTGGGTACTTCCAGGGGCAAGAGAGCCCTGGGGTC 480

Qy 1658 AGTGACACTGTGAGGCCCAACCATGCACTCCACAAAGGGGAGCATTTGGAAATGAAGGACT 1717



481 AGTGACACTCTCAGGCCCAACCATGCACTCCACAAAGGGAGGCAATTTGGAAATGAAGACT 540  
1718 AGTCTCTATGTATCAGGTTAAGACCAAGGAGAGCTGGCCAGGAGCAGCAGTGTTCACAG 1777  
541 AGTCTCTATGTATCAGGTTAAGACCAAGGAGAGCTGGCCAGGAGCAGCAGTGTTCACAG 600  
1778 CAGAGGGGAATGTAGCAACAGAGGGGCTCTTAGGCCCACTCTCCATTTCTTAGTAAAG 1837  
601 CAGAGGGGAATGTAGCAACAGAGGGGCTCTTAGGCCCACTCTCCATTTCTTAGTAAAG 660  
1838 AAGAGCATTTCTCAGACTCCAGCGGAGAGCTAGGCTTAGGCTTCAGCAACCAAGGTT 1897  
661 AAGAGCATTTCTCAGACTCCAGCGGAGAGCTAGGCTTAGGCTTCAGCAACCAAGGTT 720  
1898 CTCCTGGGACCCAAAGTTTATGGAGAAAGGCAAGACTTCATGGGAAGAGAGAGAG 1957  
721 CTCCTGGGACCCAAAGTTTATGGAGAAAGGCAAGACTTCATGGGAAGAGAGAGAG 780  
1958 GCCCTGGGTAGAAACGCTTGGTGTCTCTTTGGGCTTTAGACAAAGCGCTCATCTT 2017  
781 GGCCTGGGTAGAAACGCTTGGTGTCTCTTTGGGCTTTAGACAAAGCGCTCATCTT 839  
2018 GCCCTTACCTCTCTGATAGGCTTCAGGCTTTGCCAACACACTGTGGCTACAGGTGGAGG 2077  
840 GCCCTTACCTCTCTGATAGGCTTCAGGCTTTGCCAACACACTGTGGCTACAGGTGGAGG 898  
2078 GAAGAGACTCTCTCTCAGAGTGTATGTTCAGGAGTTCCTTTAAACCCATATGCC 2137  
899 GAAGAGACTCTCTCTCAGAGTGTATGTTCAGGAGTTCCTTTAAACCC--ATATGGC 955  
2138 CAAGAGTACTCTGAGAGGCGCTTTAAGACGGAACAAGTAATTTACCAGTTCTACT-G 2196  
956 CAAGAAATACCTTGAAGAGGCGCTTTAAGACGGAACAAGTAATTTACCAGTTCTACTGG 1015  
2197 GGGTTCTGCCACCGTCCCAAGGTGGCGAGGCTAGAGAGAGGGTCAITCTTTAAGCC 2255  
1016 GGGTTCTGCCACCGTCCCAAGGTGGCGAGGCTAGAGAGAGGGTCAITCTTTAACC 1074

RESULT 4  
LOCUS BX397021  
DEFINITION BX397021 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0DI028YM11 5-PRIME, mRNA sequence.  
ACCESSION BX397021  
VERSION BX397021.1 GI:30609314  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 956)  
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 7307.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DI028AG06\_D11486\_2&cluster=7307.r.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0DI028AG06\_D11486\_2.  
FEATURES  
source  
1..956  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI028YM11"

/tissue type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 27.1%; Score 883.4; DB 13; Length 956;  
Best Local Similarity 98.5%; Pred. No. 48-123; 9; Indels 5; Gaps 5;  
Matches 942; Conservative 0; Mismatches

1596 CAAATTTAAAGATGTGGACCAAGGCTGTGGTACTTCCAGGGCAAGAGAGCCCTGGGG 1655  
Db 2 CCNATTTAAAGATGTGGACC-AGGCTGTGGTACTTCCAGGGCAAGAGAGCCCTGGGG 60  
1656 TCAGTGACACTGTCTCAGGCCAACATGCACTCCAAAGGGGAGCATTTGGAATGAAGA 1715  
Db 61 TCAGTGACACTGTCTCAGGCCAACATGCACTCCAAAGGGGAGCATTTGGAATGAAGA 120  
1716 CTAGCTCCTATGTATCAGGTTAAGAGCAAGGAGAGCTGGCCAGGAGCAGGATTTGCAC 1775  
Db 121 CTAGCTCCTATGTATCAGGTTAAGAGCAAGGAGAGCTGGCCAGGAGCAGGATTTGCAC 180  
1776 AGCAGAGGGGAATGTAGCAACAGCAGGGGCTCTAGGCCCATCTTTCCATTTCTTAGGTA 1835  
Db 181 AGCAGAGGGGAATGTAGCAACAGCAGGGGCTCTAGGCCCATCTTTCCATTTCTTAGGTA 240  
1836 AGAAGAGCATTTCTCAGACTCCAGCGGAGAGCTAGGCTTCCAGCAACCAAGG 1895  
Db 241 AGAAGAGCATTTCTCAGACTCCAGCGGAGAGCTAGGCTTCCAGCAACCAAGG 300  
1896 TTCTCTCGGACCCAAAGTTTATGGAGAAAGGCAAGACTTCATGGGAGAGAGAGAGGA 1955  
Db 301 TTCTCTCGGACCCAAAGTTTATGGAGAAAGGCAAGACTTCATGGGAGAGAGAGAGGA 360  
1956 AGGCCCTGGGTGAAGAACGTTGTGTCTCTTTGGGCTTTAAGACAAAGCGCTCATC 2015  
Db 361 AGGCCCTGGGTGAAGAACGTTGTGTCTCTTTGGGCTTTAAGACAAAGCGCTCATC 420  
2016 TTGCCCTCTACCTCTGATAGGCTTCAGGTTTGGCCAAACACACTGTGGCTCAGGTGGA 2075  
Db 421 TTGCCCTCTACCTCTGATAGGCTTCAGGTTTGGCCAAACACACTGTGGCTCAGGTGGA 480  
2076 GCGAAGAGGACTCTCTCTCCAGAGTGTATGTTCAGGAGTTCCTTTAAACCCATATGG 2135  
Db 481 GCGAAGAGGACTCTCTCTCCAGAGTGTATGTTCAGGAGTTCCTTTAAACCCATATGG 540  
2136 CCCAAGAGTACTCTGATAGGAGGCTTTAAGAGGGAACAAGTAATTTACAGTTCTACT 2195  
Db 541 CCCAAGAGTACTCTGATAGGAGGCTTTAAGAGGGAACAAGTAATTTACAGTTCTACT 600  
2196 GGGGTTCTCGCCACCGTCCCAAGGTGGGCGAGGCTTAGGAAGAGGCTCATTTCTTAAGCC 2255  
Db 601 GGGGTTCTCGCCACCGTCCCAAGGTGGGCGAGGCTTAGGAAGAGGCTCATTTCTTAAGCC 660  
2256 ACATTTAGTGTGCACTGCGTGGCTGCAGCCAAACCAAGAACTGGTGTGTAGTATTCAT 2315  
Db 661 ACATTTAGTGTGCACTGCGTGGCTGCAGCCAAACCAAGAACTGGTGTGTAGTATTCAT 720  
2316 CAACTAAGAACCAAAATTCAGGGGCACTCATATGTGAAGGATAGAACTCAGTCTCTTAC 2375  
Db 721 CAACTAAGAACCAAAATTCAGGGGCACTCATATGTGAAGGATAGAACTCAGTCTCTTAC 780  
2376 TCCTCCAAAGAGTGGGGAAGAACCAATCAACACCTTTCTCTCTGACTTACCAAAACAG 2435  
Db 781 TCCTCCAAAGAGTGGGGAAGAACCAATCAACACCTTTCTCTCTGACTTACCAAAACAG 840  
2436 GAAACACAGCAGGAGGCTGCTCAGGACTTAGGAGCAGGATATGCTTAGATGGTGA 2495  
Db 841 G-AAACACAGCAGGAGGCTGCTCAGGACTTAGGAGCAGGATATGCTTAGATGGTGA 899  
2496 AGCAAGAGAGCAGGAGGATTTGTAATCACTGGCTTAATGAGAAAGGAGCAGCTA 2551

Db 900 AGC-AAGGAGACAGGAAGAAGTGT-AATCACTGGCT-ATGAGAAAAGAGACACTTA 952  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5  
BQ937298 LOCUS  
DEFINITION BQ937298.1 GI:22352681 linear EST 21-AUG-2002  
5', mRNA sequence.  
ACCESSION BQ937298  
VERSION BQ937298.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 957)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-x@mail.nih.gov  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LucM2664 row: m column: 12  
High quality sequence stop: 555.  
Location/Qualifiers  
1. .957  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6482195"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_40"  
/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 26.3%; Score 856; DB 13: Length 957;  
Best Local Similarity 97.4%; Pred. No. 5.2e-119;  
Matches 890; Conservative 0; Mismatches 22; Indels 2; Gaps 2;

Qy 1314 TGTGCCATGGCCCTGCCCTCCCCAGTCACCTCGAAAGTGAGCATCGAGGTAGGAGGAA 1373  
Db 1 TGTGCCATGGCCCTGCCCTCCCCAGTCACCTCGAAAGTGAGCATCGAGGTAGGAGGAA 60

Qy 1374 ACAGCAACGGGGAGTCTCTGAGCCTGGGGCTGCCCTACCTCTACCCTATCCCAGCACG 1433  
Db 61 ACGCAACCGGGAGTCTCTGAGCCTGGGGCTGCCCTACCTCTACCCTATCCCAGCACG 120

Qy 1434 AGCTTTGCCCTTGCTTGCGCTGCCCGCTGCCCTCTTTTGGGGAACGTAGCTCAGAGGCAGGT 1493  
Db 121 AGCTTTGCCCTTGCTTGCGCTGCCCGCTGCCCTCTTTTGGGGAACGTAGCTCAGAGGCAGGT 180

Qy 1494 GTTTCAGAGAAGGAACAATAATGAGGGGTGGCAGGGATAAAAAAGTCACCTCCATTCTCTA 1553  
Db 181 GC TTCAGAGAAGGAACAATAATGAGGGGTGGCAGGGATAAAAAAGTCACCTCCATTCTCTA 240

Qy 1554 CTTCCCATGAGCATGAACAATAATTTCTTCACCTGGCTCCCAAATTTAAAGATGTGGA 1613  
Db 241 CCTCCCATGAGCATGAACAATAATTTCTTCACCTGGCTCCCAAATTTAAAGATGTGGA 300

Faraday Avenue Genoscope sequence ID : CS0DI028AG06 DI1486 1.

FEATURES	source	Location/Qualifiers
	i. .915	/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="CS0D1028Yml1"
		/tissue_type="PLACENTA COT 25-NORMALIZED"
		/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
		/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN		
Query Match	26.0%; Score 847.2; DB 13; Length 915;	
Best Local Similarity	98.3%; Pred. No. 1.1e-117;	
Matches	882; Conservative 0; Mismatches 12; Indels 3; Gaps 3;	
Qy	1615	CAAGGCCCTGTGGGTACTCCAGGGCAGGAGAGCCCTGGGGTCAGTGACACTGTCAAGCC 1674
Db	22	CCAGGCCCTGTGGGTACTCCAGGGCAGGAGAGCCCTGGGGTCAGTGACACTGTCAAGCC 81
Qy	1675	AACCATGCATCTCCAAAGGGGAGCATTTGGGAATGAAGCACTAGCTTCCTATGTATCAGG 1734
Db	82	AACCATGCATCTCCAAAGGGGAGCATTTGGGAATGAAGCACTAGCTTCCTATGTATCAGG 141
Qy	1735	TTAAGACAAGGGAGAGCTGGCCAGGACAGCAGTTTTGCGACAGCAGAGGGGAATGTAGCA 1794
Db	142	TTAAGACAAGGGAGAGCTGGCCAGGACAGCAGTTTTGCGACAGCAGAGGGGAATGTAGCA 201
Qy	1795	ACAGCAGGGCTCTTAGGCCCCATCTTCCATTCTTAGGTGAAGAAGAGCATTTCTCTCAGA 1854
Db	202	ACAGCAGGGCTCTTAGGCCCCATCTTCCATTCTTAGGTGAAGAAGAGCATTTCTCTCAGA 261
Qy	1855	CTCCCAGGCGGAGCACTGAGCCTAGSCCTTCAGCAACCAAAGTTCTCTCGGACCCTAAAGT 1914
Db	262	CTCCCAGGCGGAGCACTGAGCCTAGSCCTTCAGCAACCAAAGTTCTCTCGGACCCTAAAGT 321
Qy	1915	TTATGGGAGAGGGCAAGACTTCATGGGAAGAGAGAGAGAGGCCCTGGGTAGAAAACGC 1974
Db	322	TTATGGGAGAGGGCAAGACTTCATGGGAAGAGAGAGAGAGGCCCTGGGTAGAAAACGC 381
Qy	1975	TTGGTGCTGTCTCTTTGGCCCTTTAAGCAAAAGCGCTCATCTTCGCCCTCTACCTCCCTGAT 2034
Db	382	TTGGTGCTGTCTCTTTGGCCCTTTAAGCAAAAGCGCTCATCTTCGCCCTCTACCTCCCTGAT 441
Qy	2035	AGGCTTCAGGGTTTGCCAAACACACTGTGGCTACAGGTGGAGGGAAGAGCACTCTCTTCCT 2094
Db	442	AGGCTTCAGGGTTTGCCAAACACACTGTGGCTACAGGTGGAGGGAAGAGCACTCTCTTCCT 501
Qy	2095	CCAGAGTGTATGTTCAGGAAGATTCCTTTAACCCCATATGGCCCAAGAGTAGTCGTAGG 2154
Db	502	CCAGAGTGTATGTTCAGGAAGATTCCTTTAACCCCATATGGCCCAAGAGTAGTCGTAGG 561
Qy	2155	AGGCCCTTTAAAGACGGCAACAAGTAATTTACCAAGTTCTACTGGGGTTCTTGCCCCACCGTC 2214
Db	562	AGGCCCTTTAAAGACGGCAACAAGTAATTTACCAAGTTCTACTGGGGTTCTTGCCCCACCGTC 621
Qy	2215	CCAAGTGGGGCGAGCCCTAGGAAGAGGGTCAATTTCTTAAGCCACACATTAGCTGCACCTCG 2274
Db	622	CCAAGTGGGGCGAGCCCTAGGAAGAGGGTCAATTTCTTAAGCCACACATTAGCTGCACCTCG 681
Qy	2275	TGGCTGAGCCAAAAACAAGAACTGGGTGTGTAGTATTTCATCACTAAGAACCAAAATCC 2334
Db	682	TGGCTGAGCCAAAAACAAGAACTGGGTGTGTAGTATTTCATCACTAAGAACCAAAATCC 741
Qy	2335	AGGCACTCATATGTGAAGGATAAGAACTCACCTTCTTACTCTCTCCAAAGAAAGTGGG 2394
Db	742	AGGCACTCATATGTGAAGGATAAGAACTCACCTTCTTACTCTCTTACTCTNCANNAAG-AGTGEN 800
Qy	2395	GAAAGAACCATCAAACTTTCTCTCTGATTTACCAAAACGAGGAAAAACAGCAGGAGAGGGT 2454

[illegible]

1194 GACCAAGCTCGTGGGAGGTTCCACCCCTTGATCCAGAGAAACCCCTCCACCTGCTC 1253  
Db GACCAAGCTCGTGGGAGGTTCCACCCCTTGATCCAGAGAAACCCCTCCACCTGCTC 240  
1254 GTGGGTGGCCAAAGGCTACAGGCTGCTTCTCTCTCCGCCACCCCACTGCTCCCTCA 1313  
Db GTGGGTGGCCAAAGGCTACAGGCTGCTTCTCTCTCTCCGCCACCCCACTGCTCCCTCA 300  
1314 TGTGCCATGGCCCTGCCCTCCCACTGACCTCGAAAGTGGAGCATCAGAGTAGAGGGAA 1373  
Db TGTGCCATGGCCCTGCCCTCCCACTGACCTCGAAAGTGGAGCATCAGAGTAGAGGGAA 360  
1374 ACAGCAACCGGGAGTCTCCAGGCTGGGGCTGCGCTTACCTCTACCAATCCCGGACAG 1433  
Db ACAGCAACCGGGAGTCTCCAGGCTGGGGCTGCGCTTACCTCTACCAATCCCGGACAG 420  
1434 AGCTTTCCTTGTGGCTGCGCGCTGCTCTTTGGGGAACCTGAGCTCAGAGGAGGT 1493  
Db AGCTTTCCTTGTGGCTGCGCGCTGCTCTTTGGGGAACCTGAGCTCAGAGGAGGT 480  
1494 GCTTCAGAGAGGAAACAAATAGGGGTGGCAGGATAAAAGTCACTCTCATTTCTCTA 1553  
Db GCTTCAGAGAGGAAACAAATAGGGGTGGCAGGATAAAAGTCACTCTCATTTCTCTA 540  
1554 CCTCCCATGACGATGAACAATTTCTCTCCACTGGCTCCCAATTTAAAGATGTGA 1613  
Db CCTCCCATGACGATGAACAATTTCTCTCCACTGGCTCCCAATTTAAAGATGTGA 600  
1614 CCAAGGCTGTGGTACTCCAGGGGCAAGGAGCCCTGGGGTCAGTGACACTGTGAGC 1673  
Db CCAAGGCTGTGGTACTCCAGGGGCAAGGAGCCCTGGGGTCAGTGACACTGTGAGC 660  
1674 CAACATGCTACTCCACAAGGGGAGCATTTGGAATGAAGACTAGCTTCTATGTATCAG 1733  
Db CAACATGCTACTCCACAAGGGGAGCATTTGGAATGAAGACTAGCTTCTATGTATCAG 720  
1734 GTTAAGAGCAAGGAGAGCTGGCCAGGACAGCATTTTGACAG - CAGAGGGGAATGTAG 1792  
Db GTTAAGAGCAAGGAGAGCTGGCCAGGACAGCATTTTGACAG - CAGAGGGGAGTGTAG 780  
1793 CAACAGCA - GGGCTCTTAGGCCCATCTTCATTTCTAGTGAAGAGCATTTCTC 1851  
Db CAACAGCAAGGGGCTCTTAGGCCCATCTTCATTTCTAGTGAAGAGCATTTCTC 840  
1852 AGACTCCAGCGGAGGACTCAGCTAGCC - TTGAGCAACCAAGTCTCTCTGGGACCA 1910  
Db AGACTCCAGCGGAGGAGTGAACCTAGCCCTTTCAGCAACCAAGTCTCTCTGGGAC 900  
1911 AAGTTT 1916  
Db AAGTTT 906  
RESULT 8  
BX326279 918 bp mRNA linear EST 02-MAY-2003  
LOCUS BX326279 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CSOD1028YM11 5-PRIME, mRNA sequence.  
ACCESSION BX326279  
VERSION BX326279.1 GI:30336558  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 918)  
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 7307.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAA0152B10\_CS01419\_1&cluster=7307.r.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAA0152B10\_CS01419\_1.  
Location/Qualifiers  
1. 918  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOD1028YM11"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES  
source

ORIGIN

Query Match 25.7%; Score 837.8; DB 13; Length 918;  
Best Local Similarity 97.9%; Pred. No. 2.9e-116;  
Matches 848; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 380 AGGCGTCCAGAGATCCTTATTGGGAATAAGCTGATGAGGACAGAAACGGCAGGTGG 439  
Db |||||  
Qy 52 AGGCGTCCAGAGATCCTTATTGGGAATAAGCTGATGAGGACAGAAACGGCAGGTGG 111  
Db |||||  
Qy 440 AAGAGACAAAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAAACAAGTGC 499  
Db |||||  
Qy 112 AAGAGACAAAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAAACAAGTGC 171  
Db |||||  
Qy 500 CTGACCAACCTCAACATTAAGAGTCAATTCACGCTGACAGAGCTGGTCTGCAGGC 559  
Db |||||  
Qy 172 CTGACCAACCTCAACATTAAGAGTCAATTCACGCTGACAGAGCTGGTCTGCAGGC 231  
Db |||||  
Qy 560 CCATAGGAAGGAGCTGGAAGGCTCCGATGCGTGCCAGCAATGATTTGGCACTGGCAGA 619  
Db |||||  
Qy 232 CCATAGGAAGGAGCTGGAAGGCTCCGATGCGTGCCAGCAATGATTTGGCACTGGCAGA 291  
Db |||||  
Qy 620 GCTGGAGAGGAGGAGGCGCAAAACCGAGGGCCAGCGAACTCTTCGAAAAACCTGCTGTG 679  
Db |||||  
Qy 292 GCTGGAGAGGAGGAGGCGCAAAACCGAGGGCCAGCGAACTCTTCGAAAAACCTGCTGTG 351  
Db |||||  
Qy 680 CTGAGTCTGTGTGGGACACCCACAGACACCCCTCTTCCCTCAGAGAGCCGCTGGGCA 739  
Db |||||  
Qy 352 CTGAGTCTGTGTGGGACACCCACAGACACCCCTCTTCCCTCAGAGAGCCGCTGGGCA 411  
Db |||||  
Qy 740 GACAGGGAGCGGGGCTTTGGCCCTGCTGCTCTCTCGTGTGATGACCTATTGAGTA 799  
Db |||||  
Qy 412 GACAGGGAGCGGGGCTTTGGCCCTGCTGCTGCTCTCTCGTGTGATGACCTATTGAGTA 471  
Db |||||  
Qy 800 TCAGTAGCCACTACTCCCTCCCTGCTGCGCTGAGAGGGCTCTGCTGTCATCTCAAGCAG 859  
Db |||||  
Qy 472 TCAGTAGCCACTACTCCCTCCCTGCTGCGCTGAGAGGGCTCTGCTGTCATCTCAAGCAG 531  
Db |||||  
Qy 860 CCCCTGTCCCAGCCCGTCAACCTCGAGTGGTCTTTTCAGCCTGTTTCCCGACCCACA 919  
Db |||||  
Qy 532 CCCCTGTCCCAGCCCGTCAACCTCGAGTGGTCTTTCTTACGCTGTTTCCCGACCCACA 591  
Db |||||  
Qy 920 GGCCTGTACAGCCCCCAGATGCGCGCAGAGCACTGTCTCACCATCCCGGACCCACAG 979  
Db |||||  
Qy 592 GGCCTGTACAGCCCCCAGATGCGCGCAGAGCACTGTCTCACCATCCCGGACCCACAG 651  
Db |||||  
Qy 980 ACAACAGCCAGGCTGGAGTCCAGGCCACTTTTCAGTCTCTCTTCTCCGTCATCTG 1039  
Db |||||  
Qy 652 ACAACAGCCAGGCTGGAGTCCAGGCCACTTTTCAGTCTCTCTTCTCCGTCATCTG 711  
Db |||||  
Qy 1040 CTCTTCTCTGCTTTTCTCTCTTCCCGACCTTTCTCTTCTGACCCCTCCCTCCGCTG 1099  
Db |||||  
Qy 712 CTCTTCTCTGCTTTTCTCTCTTACCCCACTTCTCTTCTCTGACCCCTCCCTCCGCTG 771  
Db |||||



Qy	1147	GTGTGAGCTCGCTCTTTCCTTCCCTTAAGCTATCAAGGGATGACCCAGGCTCGT	1206
Dd	779	GTGTGAGCTCGCTCTTTC--TTCCCTTCCAAGCTATCCAAGGGATGACCCAGGCTCGG	836
Qy	1207	GGGAGGTTCCACCCTTGGAATCAGGAAGAACCCTCCACCTGCCTCGTGGGTGGGCCAA	1266
Dd	837	GGGAGGGTTTCACTTTGGGATCCGGNAGAACCTTCC-CCCTGCTTGGGGGGGGCCNA	895
Qy	1267	AGGCTACAGGGTCTTCTTCTCTTCCCCCAC	1299
Dd	896	AGCGTACAGGTGCTTTTTTCTTTTCCCACC	928
RESULT 11			
Bil18034                                924 bp    mRNA          linear      EST 26-JUN-2001			
LOCUS                                602867301F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5015713 5',			
DEFINITION                                mRNA sequence.			
Bil18034			
Bil18034                                GI:14568935			
SOURCE                                Homo sapiens (human)			
ORGANISM                                Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
AUTHORS                                NIH-MGC http://mgc.nci.nih.gov/.			
TITLE                                National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL                                Unpublished (1999)			
COMMENT                                Contact: Robert Strausberg, Ph.D. Email: cgabbs@remail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LHCW1825 row: b column: 02 High quality sequence stop: 893.			
FEATURES			
source			
1..924			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:5015713"			
/tissue_type="small cell carcinoma"			
/cell_line="MGC3"			
/lab_host="DHIOB (phase-resistant)"			
/clone_lib="NIH_MGC_7"			
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."			
ORIGIN			
Query Match                                24.6%; Score 801.6; DB 12; Length 924;			
Best Local Similarity                                93.7%; Pred. No. 7.9e-111;			
Matches 869; Conservative                                0; Mismatches 54; Indels 4; Gaps 3;			
Qy	1504	AGGAACAATAATGAGGGTGGCAGGATAAAAGTCACTTCCATTCTTACCTCCCATGC	1563
Dd	2	AGGAACAATAATGAGGGTGGCAGGATAAAAGTCACTTCCATTCTTACCTCCCATGC	61
Qy	1564	AGCATGAACAATTTCTCTCCACTGGCTCCCAAATTTAAAGATGTGGACCAAGGCTG	1623
Dd	62	AGCATGAACAATTTCTCTCCACTGGCTCCCAAATTTAAAGATGTGGACCAAGGCTG	121
Qy	1624	TGGGTACTCCAGGGGCAAGGAGGCCCTCGGGTCACTGTGACACTGTCAAGGCCAACCATGCA	1693







```
Db 601 CTCCTATGAGCCCTTTAAAGACGGAACAAGGAATTTACCAAGTTCTACTGGGGTCTCTGC 660
Qy 2207 CCACCGTCCCAAGGTGGCGAGGCTTAGGAAGAGGGTCATTCTTTAAGCCACACATTAGCT 2266
Db 661 CCACCGTCCCAAGGTGGCGAGGCTTAGGAAGAGGGTCATTCTTTAAGCCACACATTAGCT 720
Qy 2267 GCACGTGGTGGCTGCAGCCCAAAACAAAGAACTGGGTGTGAGTATTTCATCAACTAAGAAC 2326
Db 721 GCACGTGGTGGCTGCAGCCCAAAACAAAGAACTGGGTGTGAGTATTTCATCAGCTAAGAAC 780
Qy 2327 CAATAATCAGGCGCACTCATATGT-GAAGATAAGAACC--TCACCTTCTTACTCTCTCAA 2383
Db 781 CCAATAATCCGGGGCACTCATATGTGGAAGGATAAGAACCCTTCTTCTTACTCTCTCAA 840
Qy 2384 AAACAAGTGGGG-AAAGAACCATCAAACTTTTC 2416
Db 841 AAAAAAGGGGGGAAAGAACCATCGAAACTTTTC 874

RESULT 13
BUS57215 840 bp mRNA linear EST 16-SEP-2002
LOCUS AGENCOURT 10253451 NIH MGC_109 Homo sapiens cDNA clone
DEFINITION IMAGE:6585046 5', mRNA sequence.
ACCESSION BUS57215
VERSION BUS57215.1 GI:22907511
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2796 row: f column: 22
High quality sequence start: 32
High quality sequence stop: 727.
Location/Qualifiers
1. .840
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6585046"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_109"
/notes="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; CDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 24.5%; Score 797.2; DB 13; Length 840;
Best Local Similarity 99.5%; Pred. No. 3.7e-110;
Matches 799; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2436 GAAACACGACGAGAGGGTGGCTCAGGACTTAGGACAGGGTAGCTTAGATGGTGA 2495
```

```
Db 38 GGAACACGACGAGAGGGTGGCTCAGGACTTAGGACAGGGTATAGCTTAGATGGTGA 97
Qy 2496 AGCAAAAGGAGAGAGGAGAGTGTAAATCACTGGCTAATAGAAAGAGAGACAGCTAAC 2555
Db 98 AGCAAAAGGAGAGAGGAGAGTGTAAATCACTGGCTAATAGAAAGAGAGACAGCTAAC 157
Qy 2556 TAGATGAAGCTGTGACTAGGCTGGAGTGTCTCTTGAAGATGGGACTCTTGGGTATC 2615
Db 158 TAGATGAAGCTGTGACTAGGCTGGAGTGTCTCTTGAAGATGGGACTCTTGGGTATC 217
Qy 2616 AAGACCTATGCCACATCACACTGGGGCTAGGAAAGTAGGTATGCCAGCCCTCAAGCTG 2675
Db 218 AAGACCTATGCCACATCACACTGGGGCTAGGAAAGTAGGTATGCCAGCCCTCAAGCTG 277
Qy 2676 TCTTCAGCAGGAGACTTGAAGATTATATTGGGCAGTGGCTCCCAATCTGTGGACAGTAT 2735
Db 278 TCTTCAGCAGGAGACTTGAAGATTATATTGGGCAGTGGCTCCCAATCTGTGGACAGTAT 337
Qy 2736 TTCAGCTTTCCCTGAAGATCAGGAGGGTGCATTCAATTGTCTTCTCTCTAGCCCTT 2795
Db 338 TTCAGCTTTCCCTGAAGATCAGGAGGGTGCATTCAATTGTCTTCTCTCTAGCCCTT 397
Qy 2796 CAGGAAGAGAGGACTATATTGTACTCTACCTAGGGGTCTCTGGAAGGAAACATGGA 2855
Db 398 CAGGAAGAGAGGACTATATTGTACTCTACCTAGGGGTCTCTGGAAGGAAACATGGA 457
Qy 2856 TCAGGATTTCTATAGACTGATAGGCCCTTATCCACAAGGGCCATGACTGGGAAAGTATGG 2915
Db 458 TCAGGATTTCTATAGACTGATAGGCCCTTATCCACAAGGGCCATGACTGGGAAAGTATGG 517
Qy 2916 GAGCAGAGGAGAAATGGGATTTTAGGGTGCAGTACGCTCACCTAAACTTTTGGTGGC 2975
Db 518 GAGCAGAGGAGAAATGGGATTTTAGGGTGCAGTACGCTCACCTAAACTTTTGGTGGC 577
Qy 2976 CTGGGGCATCTCTTGGGCCAGACTCTTAAAGAGGGTCTCTGGCTGTGTCTTACTCTCA 3035
Db 578 CTGGGGCATCTCTTGGGCCAGACTCTTAAAGAGGGTCTCTGGCTGTGTCTTACTCTCA 637
Qy 3036 CCACCTCTGCACCTGTCTTGGAGACTCCATCCAGCCCCCAGGACGACCTGCTCTCG 3095
Db 638 CCACCTCTGCACCTGTCTTGGAGACTCCATCCAGCCCCCAGGACGACCTGCTCTCG 697
Qy 3096 AGCTTCCACTATCTCTGTGACGGGTGAATTCGTGTACTGTGTCTGGGTCCATATAT 3155
Db 698 AGCTTCCACTATCTCTGTGACGGGTGAATTCGTGTACTGTGTCTCGTCCATATAT 757
Qy 3156 GAATTGTGAGCAGGGTTCATCTATTATAACACAGATGTTTACAAATAAGATTATTC 3215
Db 758 GAATTGTGAGCAGGGTTCATCTATTATAACACAGATGTTTACAAATAAGATTATTC 817
Qy 3216 AAACCCACCAAAAAAAAAAAAAA 3238
Db 818 AAACCAAAAAAAAAAAAAAAAAA 840

RESULT 14
BUS49866 820 bp mRNA linear EST 16-OCT-2002
LOCUS AGENCOURT 10440477 NIH MGC_109 Homo sapiens cDNA clone
DEFINITION IMAGE:6598564 5', mRNA sequence.
ACCESSION BUS49866
VERSION BUS49866.1 GI:24034829
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
```

Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2831 row: j column: 04  
High quality sequence stop: 668.  
Location/Qualifiers  
1. 820  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6598564"  
/tissue\_types="teratocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"  
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

FEATURES  
source

## ORIGIN

Query Match 24.3%; Score 792.2; DB 13; Length 820;  
Best Local Similarity 99.0%; Pred. No. 2.1e-109;  
Matches 797; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 2435 GGAACACAGCAGGAGGGTGGCTCAGGACTTAGGGACGGTATAGCTAGTGCGGA 2494  
DB 1 GGAACACAGCAGGAGGGTGGCTCAGGACTTAGGGACGGTATAGCTAGTGCGGA 60  
QY 2495 AAGCAACAGGAGAGCAGGAAGTTGTAATCACTGGCTAATGAGAAAGGACAGCTAACT 2554  
DB 61 AAGCAACAGGAGAGCAGGAAGTTGTAATCACTGGCTAATGAGAAAGGACAGCTAACT 120  
QY 2555 CTAGGATGAAGCTGTGACTAGGCTGGAGTGTCTTCTTGAAGATGGGACTCTTGGGTAT 2614  
DB 121 CTAGGATGAAGCTGTGACTAGGCTGGAGTGTCTTCTTGAAGATGGGACTCTTGGGTAT 180  
QY 2615 CAAGACCTATGCCACATCACACTGGGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCT 2674  
DB 181 CAAGACCTATGCCACATCACACTGGGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCT 240  
QY 2675 GTCTTCAGCCAGGACTTGAGAACTTATATTGGGCAGTGGCTCCAATCTGTGGACCAAGTA 2734  
DB 241 GTCTTCAGCCAGGACTTGAGAACTTATATTGGGCAGTGGCTCCAATCTGTGGACCAAGTA 300  
QY 2735 TTTCAGCTTTCCTGAAGATCAGCAGGGTGCCATTCAATTGTCTTTCTCTCTAGCCCCC 2794  
DB 301 TTTCAGCTTTCCTGAAGATCAGCAGGGTGCCATTCAATTGTCTTTCTCTCTAGCCCCC 360  
QY 2795 TCAGGAAGAGGACTATATTGTACTGTACCTAGGGTCTTGGAGGGAACATGGA 2854  
DB 361 TCAGGAAGAGGACTATATTGTACTGTACCTAGGGTCTTGGAGGGAACATGGA 420  
QY 2855 ATCAGGATTTCTATAGACTGATAGGCCCTTATCCACAGGGCCATGACTGGGAAAGGTATG 2914  
DB 421 ATCAGGATTTCTATAGACTGATAGGCCCTTATCCACAGGGCCATGACTGGGAAAGGTATG 480  
QY 2915 GGAGCAGAAGAGAAATTTAGGTTTGGGTGACGTACGCTCACCTTAAACTTTTGTGG 2974  
DB 481 GGAGCAGAAGAGAAATTTAGGTTTGGGTGACGTACGCTCACCTTAAACTTTTGTGG 540  
QY 2975 CCTGGGCGCTCTTGAAGCCCACTGTTAAGCAGGCTCTGTGGCCTGTTTACTCGTC 3034  
DB 541 CCTGGGCGCTCTTGAAGCCCACTGTTAAGCAGGCTCTGTGGCCTGTTTACTCGTC 600  
QY 3035 ACCACCTCTGCACCTGTGCTTGTAGACTTCATCCAGCCCCCAGGCACGCCACCTGCTCT 3094

DB 601 ACCACCTCTGCACCTGTGCTTGTAGACTCCATCCAGCCCGCAGCCACCTGCTCT 660  
QY 3095 GAGCTCCACTATCTCCCTGTGAGCGGTGAACCTCGTGTACTGTCTCGGTCCATATA 3154  
DB 661 GAGCTCCACTATCTCCCTGTGAGCGGTGAACCTCGTGTACTGTCTCGGTCCATATA 720  
QY 3155 TGAATTGTGAGCAGGGTTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGATTATT 3214  
DB 721 TGAATTGTGAGCAGGGTTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGATTATT 780  
QY 3215 CAAACCAACCAAAAAA 3239  
DB 781 TCAACCAACCAAAAAA 805  
RESULT 15  
BU557274  
LOCUS  
DEFINITION BU557274 863 bp mRNA linear EST 16-SEP-2002  
AGENCOURT 10253405 NIH\_MGC\_109 Homo sapiens cDNA clone  
IMAGE:6585115 5', mRNA sequence.  
ACCESSION BU557274  
VERSION BU557274.1 GI:22907570  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 863)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: gcapbs-femail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2796 row: i column: 19  
High quality sequence stop: 679.  
Location/Qualifiers  
1. 863  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6585115"  
/tissue\_types="teratocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"  
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

ORIGIN  
Query Match 24.3%; Score 792.2; DB 13; Length 863;  
Best Local Similarity 97.5%; Pred. No. 2.1e-109;  
Matches 826; Conservative 0; Mismatches 18; Indels 3; Gaps 2;  
QY 1547 TTCTCTACCTCCCATGAGCATGAACAACTTTCTCTCCACCTGGCTCCCAAAATTAAAG 1606  
DB 1 TTCTCTACCTCCCATGAGCATGAACAACTTTCTCTCCACCTGGCTCCCAAAATTAAAG 60  
QY 1607 ATGTGACCAAGCGCTGTGGGTACTCCAGGGGCAAGAGAGCCCTGGGGTCAGTGACACT 1666  
DB 61 ATGTGACCAAGCGCTGTGGGTACTCCAGGGGCAAGAGAGCCCTGGGGTCAGTGACACT 120



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:08:42 ; Search time 60 seconds  
(without alignments)

998.335 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEGRKEGPGANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	212	5	ABP62882 Human pol
2	1105	100.0	212	6	AAE29096 Human Ras
3	1105	100.0	401	4	AAU17136 Novel sig
4	1105	100.0	401	7	ADB93844 Human nov
5	1092	98.8	218	3	AAE41604 Human ORF
6	1077	97.5	212	7	ADE58127 Rat Prote
7	832	75.3	188	4	AAU17555 Novel sig
8	832	75.3	188	7	ADB94263 Human nov
9	546.5	49.5	204	4	ABE70670 Drosophil
10	545	49.3	207	7	ABE58121 Human Pro
11	545	49.3	213	4	ABE11916 Human Rab
12	540	48.9	221	5	ABP41333 Human ova
13	532	48.1	203	7	ADB83429 Human Pro
14	532	48.1	203	7	ADB58125 Human Pro
15	532	48.1	246	3	AAE58196 Lung canc
16	530.5	48.0	200	3	AAE09979 Human Rab
17	530.5	48.0	200	3	ABE19165 Amino aci
18	530.5	48.0	200	4	AAE95340 Human pro
19	530.5	48.0	218	3	AAE56993 Human pro
20	529.5	47.9	218	4	ABG07266 Novel hum
21	528.5	47.8	207	4	ABE71647 Drosophil
22	528.5	47.8	207	4	AAE67154 Amino aci
23	528.5	47.8	207	4	AAE92628 Human pro
24	528.5	47.8	207	5	ABP65204 Hypoxia-r
25	526	47.6	201	3	AAE09982 Canine Ra

#### ALIGNMENTS

##### RESULT 1

ABP62882  
ID ABP62882 standard; protein; 212 AA.

AC ABP62882;

DT 14-OCT-2002 (first entry)

DE Human polypeptide SEQ ID NO 319.

XX Human; vulnery; dermatological; neuroprotective; nontropic; cancer;  
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;  
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;  
KW burn; central nervous system disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; immune disorder;  
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.

XX Homo sapiens.

PN WO200218424-A2.

XX 07-MAR-2002.

PF 31-AUG-2001; 2001WO-US027093.

PR 01-SEP-2000; 2000US-00654935.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;

PT Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;

XX WPI; 2002-583321/62.

DR N-PSDB; ABQ93361.

XX New polynucleotide and polypeptides, useful for treatment and diagnosis  
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral  
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple  
PT sclerosis, diabetes and allergies.

XX Claim 20; SEQ ID NO 319; 284pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising one of  
CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising  
CC administering to a mammalian subject a composition comprising the protein  
CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).  
CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.  
CC (I) is useful for gene therapy of diseases and (II) can be used for

26	523	47.3	199	3	AAE09980	Human Rab
27	523	47.3	199	3	AAE09981	Human Rab
28	523	47.3	206	6	ABJ26542	Aspergill
29	523	47.3	206	6	ABJ25583	Aspergill
30	519	47.0	216	3	AAE53945	Arabidops
31	519	47.0	216	3	AAE08688	Arabidops
32	519	47.0	253	3	AAE53944	Arabidops
33	519	47.0	254	3	AAE08687	Arabidops
34	514	46.5	201	6	AAO19956	C elegans
35	513.5	46.5	215	3	AAE35215	Zea mays
36	512	46.3	190	6	AAO26373	Ras-like
37	510	46.2	218	3	AAE19220	Arabidops
38	507	45.9	216	3	AAE08006	Arabidops
39	506.5	45.8	209	4	AAE23365	Novel hum
40	501	45.3	224	3	AAE47826	Arabidops
41	501	45.3	234	3	AAE47825	Arabidops
42	501	45.3	335	4	AAE23366	Novel hum
43	479.5	43.4	256	6	AAO15989	Human Ras
44	479.5	43.4	256	7	ADA09336	Human GTP
45	478.5	43.3	201	2	AAE00919	Human Rab

CC therapeutic treatment. Diseases that may be treated include wound healing  
CC and tissue repair, burns, central nervous system disorders (e.g.  
CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral  
CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple  
CC sclerosis, diabetes and allergies. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ

Sequence 212 AA;

Query Match 100.0%; Score 1105; DB 5; Length 212;  
Best Local Similarity 100.0%; Pred. No. 1.4e-108;  
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MAKQYDLVFLRLLLIGDSGVGKTCCLLCFTDNEFHSSHSITIGVDFKMKTEVDGKIVRIQ 60  
Db 1 MAKQYDLVFLRLLLIGDSGVGKTCCLLCFTDNEFHSSHSITIGVDFKMKTEVDGKIVRIQ 60  
  
Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVDISSERTSYOHIMKWSDVDVEYAPEGVQKILIG 120  
Db 61 IWDTAGQERYQTITKQYRRAQGIPLVDISSERTSYOHIMKWSDVDVEYAPEGVQKILIG 120  
  
Qy 121 NKADBEQKRVGRGQQLAKEYGMDFYETSACTNLNIKESFTRLTTLVLAHRRKELEGL 180  
Db 121 NKADBEQKRVGRGQQLAKEYGMDFYETSACTNLNIKESFTRLTTLVLAHRRKELEGL 180  
  
Qy 181 RMRASNELALAELEEEEGKPEGPANSSKTCWC 212  
Db 181 RMRASNELALAELEEEEGKPEGPANSSKTCWC 212

RESULT 2  
AAE29096  
ID AAE29096 standard; protein; 212 AA.

AC AAE29096;

XX 24-FEB-2003 (first entry)

DE Human Ras-like protein.

XX Human; Ras-like protein; inflammation; cell proliferation; apoptosis;  
KW immunodeficiency; neurodegenerative; Alzheimer's disease; cirrhosis;  
KW Parkinson's disease; wasting disease; cachexia; myocardial infarction;  
KW osteoporosis; atherosclerosis; glomerulonephritis; Crohn's disease;  
KW irritable bowel syndrome; multiple sclerosis; osteoarthritis; allergy;  
KW pancreatitis; autoimmune thyroiditis; rheumatoid arthritis; osteopathic;  
KW Sjogren's syndrome; infection; transgenic; gene therapy; neutropenic; gout;  
KW neuroprotective; vasotropic; cytostatic; dermatological; nephrotropic;  
KW ophthalmological; tranquilizer; cancer; stroke; Grave's disease; AIDS;  
KW asthma; anaemia; drug screening.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Binding-site 15..22 /note= "ATP/GTP binding site motif A"

FT Modified-site 18..23 /note= "N-myristoylation site"

FT Modified-site 29..32 /note= "Casein kinase II phosphorylation site"

FT Modified-site 92..94 /note= "Protein kinase C phosphorylation site"

FT Modified-site 101..109 /note= "Tyrosine kinase phosphorylation site"

FT Modified-site 104..107 /note= "Casein kinase II phosphorylation site"

FT Modified-site 136..141 /note= "N-myristoylation site"

FT Modified-site 205..208 /note= "N-glycosylation site"

FT Modified-site 206..208

FT /note= "Protein kinase C phosphorylation site"

XX WO200277193-A2.

XX 03-OCT-2002.

XX 27-MAR-2002; 2002WO-US009328.

XX 27-MAR-2001; 2001US-00817198.

XX (PEKE ) PE CORP.

XX Gan W, Ye J, Di Francesco V, Beasley EM;

XX WPI; 2003-018913/01.

XX N-PSDB; AAD47168, AAD47619.

XX New isolated human Ras-like protein polypeptide, useful for diagnosing,  
PT treating or preventing inflammation and disorders associated with cell  
PT proliferation and apoptosis, e.g. AIDS, cancer, allergies, anemia, asthma  
PT or stroke.

PS Claim 1; Page 74; 82pp; English.

XX The invention relates to human Ras-like protein and its corresponding  
CC nucleic acid. The Ras-like protein and DNA is useful in the development  
CC of human therapeutics and diagnostic compositions. They are useful in the  
CC diagnosis, prevention and treatment of inflammation and disorders  
CC associated with cell proliferation and apoptosis, e.g. AIDS and other  
CC infectious or genetic immunodeficiencies, neurodegenerative disease e.g.  
CC Alzheimer's disease, Parkinson's disease, wasting diseases e.g. cachexia,  
CC ischaemic injuries e.g. myocardial infarction, stroke or reperfusion  
CC injury, toxin-induced diseases such as alcohol-induced liver damage or  
CC cirrhosis, osteoporosis or cancer. They are also used to treat disorders  
CC associated with inflammation including allergies, atopic dermatitis,  
CC atherosclerosis, asthma, anaemia, Crohn's disease, diabetes mellitus,  
CC lupus erythematosus, multiple sclerosis, osteoarthritis, pancreatitis,  
CC autoimmune thyroiditis, rheumatoid arthritis, Sjogren's syndrome,  
CC uveitis, trauma, or viral, bacterial, fungal, parasitic, protozoal or  
CC helminthic infections. The antibodies of the invention are useful in  
CC pharmacogenomic analysis or for tissue typing. The transgenic animals are  
CC useful for studying the function of a Ras-like protein, and identifying  
CC and evaluating modulators of its activity. Ras-like protein is used in  
CC drug screening assays and its DNA is used in gene therapy. The present  
CC sequence is human Ras-like protein

XX Sequence 212 AA;

Query Match 100.0%; Score 1105; DB 6; Length 212;  
Best Local Similarity 100.0%; Pred. No. 1.4e-108;  
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAKQYDLVFLRLLLIGDSGVGKTCCLLCFTDNEFHSSHSITIGVDFKMKTEVDGKIVRIQ 60  
Db 1 MAKQYDLVFLRLLLIGDSGVGKTCCLLCFTDNEFHSSHSITIGVDFKMKTEVDGKIVRIQ 60  
  
Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVDISSERTSYOHIMKWSDVDVEYAPEGVQKILIG 120  
Db 61 IWDTAGQERYQTITKQYRRAQGIPLVDISSERTSYOHIMKWSDVDVEYAPEGVQKILIG 120  
  
Qy 121 NKADBEQKRVGRGQQLAKEYGMDFYETSACTNLNIKESFTRLTTLVLAHRRKELEGL 180  
Db 121 NKADBEQKRVGRGQQLAKEYGMDFYETSACTNLNIKESFTRLTTLVLAHRRKELEGL 180  
  
Qy 181 RMRASNELALAELEEEEGKPEGPANSSKTCWC 212  
Db 181 RMRASNELALAELEEEEGKPEGPANSSKTCWC 212

RESULT 3  
AAU17136  
ID AAU17136 standard; protein; 401 AA.

XX AAU17136;  
XX  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Novel signal transduction pathway protein, Seq ID 701.  
XX  
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX  
XX W0200154733-A1.  
XX  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US001312.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184684P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.



PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465460/50.  
XX N-PSDB; AAS27053.  
XX  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
XX prognosing disorders related to the proteins, including cancers, immune  
XX disorders and neuronal disorders.  
XX  
XX Claim 1; SEQ ID NO 701; 880pp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
XX diagnosing, preventing and treating diseases including immune system  
XX disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
XX disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
XX transplant rejections and graft versus host disease, infectious diseases  
XX (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
XX other blood-related disorders (sickle cell anaemia), myeloproliferative  
XX disorders, primary haematopoietic disorders, hyperproliferative disorders  
XX (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.  
XX Alzheimer's disease, Parkinson's disease), chromosomal abnormalities  
XX (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.  
XX glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),  
XX respiratory disorders, dermatological disorders, in wound healing,  
XX epithelial cell proliferation, endocrine disorders (e.g. Addison's  
XX disease), reproductive system disorders, gastrointestinal disorder  
XX (inflammatory disorders), liver disorders (cirrhosis), as stimulators of  
XX B-cell responsiveness to pathogens, activators of T-cells, to induce  
XX higher affinity antibodies, and as a means to induce tumour proliferation  
XX in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAU17059-  
XX AAU17693 represent novel signal transduction pathway protein, amino acid  
XX sequences of the invention

Query Match 100.0%; Score 1105; DB 4; Length 401;  
Best Local Similarity 100.0%; Pred. No. 3.5e-108;  
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVKRIQ 60  
Db 27 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVKRIQ 86  
QY 61 IWDTAGQERYQITKQYRRAQGI FLVYDYSERSYQHIMKMWSDVDEYAPGVOKILIG 120  
Db 87 IWDTAGQERYQITKQYRRAQGI FLVYDYSERSYQHIMKMWSDVDEYAPGVOKILIG 146  
QY 121 NKADEBQKQVREGQOQQLAKEYGMDFYETSACTNLNLIKESFRTLTELVLQAKRKEGL 180  
Db 147 NKADEBQKQVREGQOQQLAKEYGMDFYETSACTNLNLIKESFRTLTELVLQAKRKEGL 206  
QY 181 RMRASNELALAELEEGKPEGPANSSKTCWC 212  
Db 207 RMRASNELALAELEEGKPEGPANSSKTCWC 238

RESULT 4  
ADB93844  
ID ADB93844 standard; protein; 401 AA.  
XX  
XX ADB93844;  
XX  
XX 04-DEC-2003 (first entry)  
XX  
XX Human novel protein #78.  
XX  
XX human; autoimmune disease; Parkinson's disease; silicosis;  
XX gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;  
XX immunosuppressive agent; adjuvant; enhance immune response;  
XX higher affinity antibody induction;  
XX increased serum immunoglobulin concentration.  
XX  
XX Homo sapiens.  
XX  
XX US2002168711-A1.  
XX  
XX 14-NOV-2002.  
XX  
XX 17-JAN-2001; 2001US-00764868.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225268P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225447P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 22-AUG-2000; 2000US-0226868P.  
XX 30-AUG-2000; 2000US-0228924P.  
XX 01-SEP-2000; 2000US-0229287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229509P.  
XX 05-SEP-2000; 2000US-0229513P.  
XX 08-SEP-2000; 2000US-0231413P.  
XX 21-SEP-2000; 2000US-0234223P.  
XX 21-SEP-2000; 2000US-0234274P.  
XX 25-SEP-2000; 2000US-0234997P.  
XX 27-SEP-2000; 2000US-0235834P.  
XX 29-SEP-2000; 2000US-0236327P.  
XX 29-SEP-2000; 2000US-0236367P.  
XX 29-SEP-2000; 2000US-0236388P.  
XX 29-SEP-2000; 2000US-0236369P.  
XX 29-SEP-2000; 2000US-0236370P.  
XX 02-OCT-2000; 2000US-0236802P.  
XX 02-OCT-2000; 2000US-0237037P.  
XX 02-OCT-2000; 2000US-0237038P.  
XX 02-OCT-2000; 2000US-0237039P.  
XX 02-OCT-2000; 2000US-0237040P.  
XX 13-OCT-2000; 2000US-0239935P.  
XX 20-OCT-2000; 2000US-0240960P.  
XX 20-OCT-2000; 2000US-0241785P.  
XX 20-OCT-2000; 2000US-0241809P.  
XX 01-NOV-2000; 2000US-0244617P.  
XX 17-NOV-2000; 2000US-0249299P.

PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251869P.  
XX  
PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI; 2003-71985/68.  
DR N-PSDB; ADB93221.  
XX  
PT New isolated polypeptide useful for diagnosing and treating  
PT immunosuppressive conditions such as autoimmune disease and Parkinson's  
PT disease.  
XX  
PS Claim 11; SEQ ID NO 701; 345pp; English.  
XX  
XX The invention relates to an isolated polypeptide. The polypeptide is  
CC useful for diagnosing a pathological condition or a susceptibility to a  
CC pathological condition in a subject, by determining the presence or  
CC amount of expression of the polypeptide in a biological sample and  
CC diagnosing a pathological condition or a susceptibility to a pathological  
CC condition based on the presence or amount of expression of the  
CC polypeptide. The polypeptide is also useful for identifying a binding  
CC partner to the polypeptide, which involves contacting the polypeptide  
CC with a binding partner and determining whether the binding partner  
CC effects an activity of the polypeptide. The polypeptide or the nucleic  
CC acid encoding the polypeptide is useful for preventing, treating, or  
CC ameliorating a medical condition, which involves administering the  
CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid  
CC is useful for diagnosing a pathological condition or a susceptibility to  
CC a pathological condition in a subject, which involves determining the  
CC presence or absence of a mutation in the nucleic acid, and diagnosing a  
CC pathological condition or susceptibility to a pathological condition  
CC based on the presence or absence of the mutation. The polypeptide, the  
CC nucleic acid and an antibody to the polypeptide are useful for treating  
CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal  
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,  
CC the nucleic acid and the antibody are useful as immunosuppressive agents,  
CC as adjuvants to enhance immune responses, and as agents to induce higher  
CC affinity antibodies and increase serum immunoglobulin concentrations. The  
CC present sequence represents the amino acid sequence of a novel human  
CC protein. Note: The sequence data for this patent did not form part of the  
CC printed specification but was obtained in electronic format direct from  
CC USPTO at seqdata.uspto.gov/sequence.html?DocID=20020168711.  
XX  
XX Sequence 401 AA;

Query Match 100.0%; Score 1105; DB 7; Length 401;  
Best Local Similarity 100.0%; Pred. No. 3.5e-108;  
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLIGDSVGKTCLLCRFTDNEFHSSHSITIGVDFPMKTIKVRQ 60  
Db 27 MAKQYDVLFRLLIGDSVGKTCLLCRFTDNEFHSSHSITIGVDFPMKTIKVRQ 86  
Qy 61 IWDTAGQRYQITTKQYRRAGQIFLVYDTSERSYQHMKWSDVDYAEQVKILIG 120  
Db 87 IWDTAGQRYQITTKQYRRAGQIFLVYDTSERSYQHMKWSDVDYAEQVKILIG 146  
Qy 121 NKABEQKRGVREGQQLAKYGMDFYETASCTNLNFKESFTRLTELVLQAHKLEGL 180  
Db 147 NKABEQKRGVREGQQLAKYGMDFYETASCTNLNFKESFTRLTELVLQAHKLEGL 206  
Qy 181 RMRASNELALAELEEEKPEGPANSSKTCWC 212  
Db 207 RMRASNELALAELEEEKPEGPANSSKTCWC 238

ID AAB41604 standard; protein; 218 AA.  
XX AAB41604;  
DT 08-FEB-2001 (first entry)  
DE Human ORFX ORF1368 polypeptide sequence SEQ ID NO:2736.  
XX  
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
KW vulnary; antiparotiatric; antiparotiatric; antiparotiatric; antiparotiatric;  
KW anticonvulsant; osteopathic; antiparotiatric; antiparotiatric; antiparotiatric;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiparotiatric;  
KW antiviral; antiparotiatric; antiparotiatric; antiparotiatric; antiparotiatric;  
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive.  
XX Homo sapiens.  
OS  
XX WO200058473-A2.  
PN  
XX 05-OCT-2000.  
PD  
XX 31-MAR-2000; 2000WO-US008621.  
PF  
XX 31-MAR-1999; 99US-0127607P.  
PR 02-APR-1999; 99US-0127636P.  
PR 05-APR-1999; 99US-0127728P.  
PR 30-MAR-2000; 2000US-00540763.  
XX (CURA-) CURAGEN CORP.  
PA  
PI Shimkets RA, Leach M;  
XX  
XX WPI; 2000-602362/57.  
DR N-PSDB; AAC75813.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease.  
XX  
PS Claim 11; Page 1979-1980; 5507pp; English.  
XX  
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
CC antiparotiatric; antiparotiatric; antiparotiatric; antiparotiatric;  
CC anticonvulsant; antiparotiatric; antiparotiatric; antiparotiatric;  
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antiparotiatric; antiparotiatric;  
CC antiviral; antiparotiatric; antiparotiatric; antiparotiatric; antiparotiatric;  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
XX Sequence 218 AA;

Best Local Similarity 97.2%; Pred. No. 3.4e-107;  
Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;  
QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKIVRIQ 60  
DB 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKIVRIQ 60  
QY 61 IWDTAGOERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDYAPGQVKILIG 120  
DB 61 IWDTAGOERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDYAPGQVKILIG 120  
QY 121 NKADEBKQKRGVGGQ-----OLAKYGMDFYETSACTNLNLIKESFTRLTELVLQAH 174  
DB 121 NKADEBKQKRGVGGQKQKPSQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQAH 180  
QY 175 KELEGLMRASNELALAELEBEKGPEGPANSSKTCWC 212  
DB 181 KELEGLMRASNELALAELEBEKGPEGPANSSKTCWC 218  
RESULT 6  
ID ADE58127 standard; protein; 212 AA.  
AC ADE58127;  
XX 29-JAN-2004 (first entry)  
DT Rat Protein AAA41995, SEQ ID NO 3998.  
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX Rattus norvegicus.  
OS WO2003016475-A2.  
PN 27-FEB-2003.  
XX 14-AUG-2002; 2002WO-US025765.  
PF 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX (GEO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
DR GENBANK; AAA41995.  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX Claim 1; Page; 1017pp; English.  
PS The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the

activity in an animal of one or more of the polypeptides given in the  
specification, a method for identifying a compound useful in treating  
pain and a pharmaceutical composition comprising the one or more  
polypeptides or their antibodies. The polynucleotide or the compound that  
modulates its activity is useful for preparing a medicament for treating  
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
therapy). The sequence presented is a rat protein (shown in Table 2 of  
the specification) which is differentially expressed during pain. Note:  
The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 212 AA;  
QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKIVRIQ 60  
DB 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKIVRIQ 60  
QY 61 IWDTAGOERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDYAPGQVKILIG 120  
DB 61 IWDTAGOERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDYAPGQVKILIG 120  
QY 121 NKADEBKQKRGVGGQKQKPSQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQAH 180  
DB 121 NKADEBKQKRGVGGQKQKPSQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQAH 180  
QY 181 RMCASNELALAELEBEKGPEGPANSSKTCWC 212  
DB 181 RMCASNELALAELEBEKGPEGPANSSKTCWC 212  
RESULT 7  
ID AAU17555 standard; protein; 188 AA.  
AC AAU17555;  
XX 07-NOV-2001 (first entry)  
DT Novel signal transduction pathway protein, Seq ID 1120.  
DE Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
XX acquired immune deficiency syndrome.  
OS Homo sapiens.  
XX WO200154733-A1.  
XX 02-AUG-2001.  
PF 17-JAN-2001; 2001WO-US001312.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-APR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229533P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249257P.  
PR 17-NOV-2000; 2000US-0249259P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254037P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465460/50.  
XX N-PSDB; AAS27472.  
XX  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
XX prognosing disorders related to the proteins, including cancers, immune  
XX disorders and neuronal disorders.  
PS Claim 1; SEQ ID NO 1120; 880pp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders).

disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ transplant rejections and graft versus host disease, infectious diseases (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and other blood-related disorders (sickle cell anaemia), myeloproliferative disorders, primary haematopoietic disorders, hyperproliferative disorders (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmia), respiratory disorders, dermatological disorders, in wound healing, epithelial cell proliferation, endocrine disorders (e.g. Addison's disease), reproductive system disorders, gastrointestinal disorder (inflammatory disorders), liver disorders (cirrhosis), as stimulators of B-cell responsiveness to pathogens, activators of T-cells, to induce higher affinity antibodies, and as a means to induce tumour proliferation in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction pathway protein, amino acid sequences of the invention

Query Match 75.3%; Score 832; DB 4; Length 188;  
Best Local Similarity 97.6%; Pred. No. 1.1e-79;  
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLLCFTDNEFHSSHSITGVDFKMTIEVDGKVRIQ 60  
DB 24 MAKQYDVLFRLLIGDSGVGKTCCLLCFTDNEFHSSHSITGVDFKMTIEVDGKVRIQ 83

QY 61 IWDTAGERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGQVKILIG 120  
DB 84 IWDTAGERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGQVKILIG 143

QY 121 NKADEEQRQVGRGQQLAKYEGMDPYETSACTNLIKESFTR 164  
DB 144 NKADEEQRQVGRGQQLAKYEGMDPYETSACTNLIKESFTR 187

RESULT 8  
ADB94263  
ID ADB94263 standard; protein; 188 AA.

XX AC ADB94263;  
XX DT 04-DEC-2003 (first entry)  
XX DE Human novel protein #497.  
XX KW human; autoimmune disease; Parkinson's disease; silicosis;  
KW gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;  
KW immunosuppressive agent; adjuvant; enhance immune response;  
KW higher affinity antibody induction;  
KW increased serum immunoglobulin concentration.

XX OS Homo sapiens.

XX PN US2002168711-A1.

XX PD 14-NOV-2002.

XX PF 17-JAN-2001; 2001US-00764868.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239355P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.

XX (RUBE/) RUBEN S M.

XX (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2003-719985/68.

XX N-PSDB; ADB93640.

XX New isolated polypeptide useful for diagnosing and treating immunosuppressive conditions such as autoimmune disease and Parkinson's disease.

XX Claim 11; SEQ ID NO 1120; 345pp; English.

XX The invention relates to an isolated polypeptide. The polypeptide is useful for diagnosing a pathological condition or a susceptibility to a pathological condition in a subject, by determining the presence or amount of expression of the polypeptide in a biological sample and diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide. The polypeptide is also useful for identifying a binding partner to the polypeptide, which involves contacting the polypeptide with a binding partner and determining whether the binding partner effects an activity of the polypeptide. The polypeptide or the nucleic acid encoding the polypeptide is useful for preventing, treating, or ameliorating a medical condition, which involves administering the polypeptide or the nucleic acid to a mammalian subject. The nucleic acid is useful for diagnosing a pathological condition or a susceptibility to a pathological condition in a subject, which involves determining the presence or absence of a mutation in the nucleic acid, and diagnosing a pathological condition or susceptibility to a pathological condition based on the presence or absence of the mutation. The polypeptide, the nucleic acid and an antibody to the polypeptide are useful for treating

CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal  
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,  
CC the nucleic acid and the antibody are useful as immunosuppressive agents,  
CC as adjuvants to enhance immune responses, and as agents to induce higher  
CC affinity antibodies and increase serum immunoglobulin concentrations. The  
CC present sequence represents the amino acid sequence of a novel human  
CC protein. Note: The sequence data for this patent did not form part of the  
CC printed specification but was obtained in electronic format direct from  
CC USPTO at [seqdata.uspto.gov/sequence.html?DocID=20020168711](http://seqdata.uspto.gov/sequence.html?DocID=20020168711).

XX SQ Sequence 188 AA;  
Query Match 75.3%; Score 832; DB 7; Length 188;  
Best Local Similarity 97.6%; Pred. No. 1.1e-79;  
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MAKQYDVLFRLLLLIGDSGVGKTCILCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVIQ 60  
DB 24 MAKQYDVLFRLLLLIGDSGVGKTCILCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVIQ 83  
QY 61 IWDTAGQERYQITTKQYRRAGQIFLVYDIISSERSYQHIMKWSDVDVEYAPGVQKILIG 120  
DB 84 IWDTAGQERYQITTKQYRRAGQIFLVYDIISSERSYQHIMKWSDVDVEYAPGVQKILIG 143  
QY 121 NKADEQKQVGRGQQLAKYGMDFYETSACTNLNIKESFTR 164  
DB 144 NKADEQKQVGRGQQLAKYGMDFYETSACTNLNIKESFTR 187

RESULT 9  
ABB70670  
ID ABB70670 standard; protein; 204 AA.  
XX AC ABB70670;  
XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster polypeptide SEQ ID NO 38802.  
XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX KW pharmaceutical.  
XX OS Drosophila melanogaster.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PF 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
XX PR 11-JUL-2000; 2000US-00614150.  
XX PA (PEKE ) PE CORP NY.  
XX PI Venter JC, Adams M, Li PWD, Myers EW;  
XX DR WPI; 2001-656860/75.  
XX DR N-PSDB; ABL14773.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX PS Disclosure; SEQ ID NO 38802; 21pp + Sequence Listing; English.  
XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
XX CC useful in developmental biology and in elucidating cell signalling and  
XX CC cell-cell interactions in higher eukaryotes for the development of  
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)  
XX SQ Sequence 204 AA;  
Query Match 49.5%; Score 546.5; DB 4; Length 204;  
Best Local Similarity 49.8%; Pred. No. 2.3e-49;  
Matches 105; Conservative 47; Mismatches 50; Indels 9; Gaps 3;  
QY 1 MAKQYDVLFRLLLLIGDSGVGKTCILCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVI 59  
DB 1 MAKQYDVLFRLLLLIGDSGVGKTCILCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVI 60  
QY 60 QIWDTAGQERYQITTKQYRRAGQIFLVYDIISSERSYQHIMKWSDVDVEYAPGVQKILI 119  
DB 61 QIWDTAGQERYQITTKQYRRAGQIFLVYDIISSERSYQHIMKWSDVDVEYAPGVQKILI 120  
QY 120 GNKADEQKQVGRGQQLAKYGMDFYETSACTNLNIKESFTRLTTELVLQAHKELEG 179  
DB 121 GNKCDMTDKRVVKNKERGEATAREHGIRFEMETSAKSNINIERAFCELAAILD---KTSG 176  
QY 180 LRWRASNELALAELEEEGKPEGPANSSKTC 210  
DB 177 RESAENQERVIIIDRRNQEKAP----GYSKCC 203

RESULT 10  
ADE58121  
ID ADE58121 standard; protein; 207 AA.  
XX AC ADE58121;  
XX DT 29-JAN-2004 (first entry)  
XX DE Human Protein P24407, SEQ ID NO 3992.  
XX KW Human; pain; neuronal tissue; gene therapy;  
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
XX KW spared nerve injury; SNI; Chung.  
XX OS Homo sapiens.  
XX PN WO2003016475-A2.  
XX PD 27-FEB-2003.  
XX PF 14-AUG-2002; 2002WO-US025765.  
XX PR 14-AUG-2001; 2001US-0312147P.  
XX PR 01-NOV-2001; 2001US-0346382P.  
XX PR 26-NOV-2001; 2001US-0333347P.  
XX PA (GEHO ) GEN HOSPITAL CORP.  
XX PA (FARB ) BAYER AG.  
XX PI Woolf C, D'urso D, Befort K, Costigan M;  
XX DR WPI; 2003-268312/26.  
XX DR GENBANK; P24407.

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX PS Claim 1; Page; 1017pp; English.  
XX CC The invention discloses a composition comprising two or more isolated rat  
XX CC or human polynucleotides or a polynucleotide which represents a fragment,  
XX CC derivative or allelic variation of the nucleic acid sequence. Also  
XX CC claimed are a vector comprising the novel polynucleotide, a host cell  
XX CC comprising the vector, a method for identifying a nucleotide sequence  
XX CC which is differentially regulated in an animal subjected to pain and a  
XX CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 207 AA;

Query Match 49.3%; Score 545; DB 7; Length 207;  
Best Local Similarity 52.9%; Pred. No. 3.4e-49;  
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;  
Qy 1 MAKQYDLVFRLLIGDSGVGKTCCLLCFTDNEPHSHSTIGVDFKMTTEVDGKIRIQ 60  
Db 1 MAKTYDLFKLLIGDSGVGKTCVLFPSDFAFNSTFTIGIDFKIRTIELDGKRIKLQ 60  
Qy 61 IWDTAGQERYOTITKQYVRAAGIPLVYDYSSESYOHIMKWSVDVYAPGEGVKILIG 120  
Db 61 IWDTAGQERFRTITTYAYRGANGIMLVYDITNEKSFNIRNWIKNIEHSASADVEKMILG 120  
Qy 121 NKADDEOKRVGRQGOQLAKEYGMDFYETSACTNLNIKESFTLTLVLVLAHQHKELEGL 180  
Db 121 NKCDVNDKQVSKERGEKALDYGKFMETSAKANINVENAFFTLARDIKAKMDKLEGN 180  
Qy 181 RWRASNE 187  
Db 181 SPOGSNQ 187

RESULT 11

ID ABB11916 standard; peptide; 213 AA.

AC ABB11916;

DT 11-JAN-2002 (first entry)

XX Human rab8 homologue, SEQ ID NO:2286.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; antiinflammatory;  
KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;  
KW cytosstatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
KW antifungal; vulneryary; antiulcer.

OS Homo sapiens.

XX WO200157188-A2.

PN 09-AUG-2001.

XX .

XX 05-FEB-2001; 2001WO-US003800.  
XX 03-FEB-2000; 2000US-00496914.  
XX 27-APR-2000; 2000US-00560875.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
PI WPI; 2001-457740/49.  
XX N-PSDB; ABA09160.  
XX Human proteins and DNA encoding sequences useful for preventing, treating  
XX or ameliorating a medical condition in a mammalian subject e.g. arthritis  
XX and cancer.  
XX Claim 20; Page 276; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
XX invention also relates to vectors and recombinant host cells comprising a  
XX nucleotide of the invention, methods of producing the novel polypeptides,  
XX antibodies against the polypeptides, methods of detecting the nucleotides  
XX or polypeptides in a sample, and methods of identifying compounds which  
XX bind to polypeptides of the invention. Although novel, many of the  
XX polypeptides of the invention have homology to known proteins, and hence  
XX giving an insight into their probable biological activities, and hence  
XX potential therapeutic applications. The polypeptides of the invention may  
XX have various activities, including cytokine, cell proliferation or cell  
XX differentiation activities; stem cell growth factor activity;  
XX haematopoiesis regulatory activity; tissue growth activity;  
XX immunomodulatory activity; activin- or inhibin-related activities;  
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or  
XX thrombolytic activities; receptor or ligand activities; or may be  
XX involved in oncogenesis; cancer cell proliferation or metastasis.  
XX Depending on their biological activities, polypeptides and nucleotides of  
XX the invention are useful for preventing, treating or ameliorating medical  
XX conditions, e.g., by protein or gene therapy. Such conditions include  
XX cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
XX proliferative retinopathy, atherosclerosis, coronary heart disease,  
XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
XX vascular growth. Polypeptides involved with tissue regeneration and  
XX repair (or nucleic acids encoding them) may be used to promote wound  
XX healing (e.g., of burns, incisions and ulcers), while those with  
XX immunomodulatory activities may be used in the treatment of viral,  
XX bacterial and fungal infections in addition to immune disorders.  
XX Polypeptides with growth factor activity may be used in cell cultures to  
XX promote cell growth. For example, such polypeptides may be used to  
XX manipulate stem cells in culture to give rise to neuroepithelial cells  
XX that can be used to augment or replace cells damaged by illness,  
XX autoimmune disease or accidental damage. The polypeptides and nucleotides  
XX may also be used in the diagnosis of the above conditions, and in drug  
XX screening techniques. The present sequence represents a novel human  
XX polypeptide of the invention

XX Sequence 213 AA;

Query Match 49.3%; Score 545; DB 4; Length 213;  
Best Local Similarity 52.9%; Pred. No. 3.6e-49;  
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MAKQYDLVFRLLIGDSGVGKTCCLLCFTDNEPHSHSTIGVDFKMTTEVDGKIRIQ 60  
Db 7 MAKTYDLFKLLIGDSGVGKTCVLFPSDFAFNSTFTIGIDFKIRTIELDGKRIKLQ 66  
Qy 61 IWDTAGQERYOTITKQYVRAAGIPLVYDYSSESYOHIMKWSVDVYAPGEGVKILIG 120  
Db 67 IWDTAGQERFRTITTYAYRGANGIMLVYDITNEKSFNIRNWIKNIEHSASADVEKMILG 126  
Qy 121 NKADDEOKRVGRQGOQLAKEYGMDFYETSACTNLNIKESFTLTLVLVLAHQHKELEGL 180



Db 127 NKDVNDKRVSKERGEKALDYGKFMETSAKANINVENAFTLARDIKAKMDKKLEGN 186

Qy 181 RMRASNE 187

Db 187 SPQGSNQ 193

RESULT 12

ID ABP41333 standard; protein; 221 AA.

XX AC ABP41333;

XX DT 23-AUG-2002 (first entry)

XX DE Human ovarian antigen HCGMA67, SEQ ID NO:2465.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KW KW ovarian cancer; breast cancer; reproductive system disorder; infertility;

KW pregnancy disorder; anovulation; polycystic ovary syndrome; PCOS;

KW ovarian cyst; dysmenorrhoea; endocrine disorder; infection;

KW inflammatory condition; immune disorder; blood disorder;

KW cardiovascular disorder; respiratory disorder; neurological disorder;

KW gastrointestinal disorder; urinary system disorder; drug screening;

KW gene therapy; chromosome mapping; forensic analysis;

KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

KW antiinflammatory; gynaecological; reproductive.

XX OS Homo sapiens.

XX FN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US018569.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-ESDB; ABQ54410.

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,

PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian

PT cancer), immune disorders, cardiovascular disorders and neurological

PT diseases.

XX PS Claim 11; SEQ ID NO 2465; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-

CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also

CC encompasses polypeptides 90% identical and polynucleotides 95% identical

CC to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human ovarian antigen

CC polynucleotides, antibodies against human ovarian antigens, and the use

CC of ovarian antigen polynucleotides and polypeptides in diagnosing,

CC treating, prognosing or preventing various ovary and/or breast-related

CC disorders. Such conditions include ovarian cancer and breast cancer, and

CC metastatic tumours of ovarian or breast origin, reproductive system

CC disorders (e.g., infertility, disorders of pregnancy, anovulation,

CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine

CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic

CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and

CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),

CC blood-related disorders (e.g., anaemia), cardiovascular disorders,

CC respiratory disorders, neurological disorders, gastrointestinal disorders

CC and urinary system disorders. Ovarian antigen polypeptides and

CC polynucleotides may also be used in screening for compounds which

CC modulate ovarian antigen expression or activity. The polynucleotides may

CC further be used for gene therapy, chromosome mapping, in the

CC identification of individuals and in forensic analysis, and the

CC polypeptides may be used as food additives or to prepare antibodies

CC useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention. Note: The

CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 221 AA;

Query Match 48.9%; Score 540; DB 5; Length 221;

Best Local Similarity 52.7%; Pred. No. 1.3e-48;

Matches 98; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Qy 2 AKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITGVDKFMKTIYVDGKIVRIQI 61

Db 16 AKTYDYLFKLLIGDSGVGKTCVLFSEDAFNFTSTIGIDFKIRTIELDGKRIKLIQI 75

Qy 62 WDTAGQERYQTITKQYVRRAGQIFLVYDISSERSYQHIMKWVSDVDEYAPGVOKILIGN 121

Db 76 WDTAGQERFRITITAYYRGAMGIMLVYDITNEKSFDIRNWRNIEEHASADVEKMILGN 135

Qy 122 KADBEQKRVGRGQGOOLAKKEYGMDFYETSACTNINIKESPTRLTELVLQHRKELEGLR 181

Db 136 KCDVNDKRVSKERGEKALDYGKFMETSAKANINVENAFTLARDIKAKMDKKLEGN 195

Qy 182 MRASNE 187

Db 196 PQGSNQ 201

RESULT 13

ADE83429

ID ADE83429 standard; protein; 203 AA.

XX AC ADE83429;

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein P51153, SEQ ID NO 11024.

XX KW Human; pain; neuronal tissue; gene therapy;

KW KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX FN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO ) GEN HOSPITAL CORP.

XX PA (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P51153.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,

```
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 203 AA;

Query Match 48.1%; Score 532; DB 7; Length 203;
Best Local Similarity 47.6%; Pred. No. 8e-48;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDFPKMTIEVDGKVRIO 60
DB 1 MAKAYDHLFKLLIGDSGVGKTCLLIRFAEDNFNTYISTIGIDFKIRTVDIEGKKIKLQ 60

QY 61 IWDTAGOERYOTITKQYVRAAGIFLVYDISSERSYOHIMKWVSDVDEYAPEGVOKILIG 120
DB 61 VMDTAGOERFKTITTAIRYGAMGILVYDITDEKSFENIQNMWKSIVENASAGVERILLG 120

QY 121 NKADDEQKROVREGQOOLAKKEYGMDFYETSACTNLNIKESFTRITELVLQAHKKELEGL 180
DB 121 NKCDMEAKRKVKQEQADKLAREHGIRFFETSASKSMNVDEAFSSLDIL-----LKSG 174

QY 181 RMRASNELALAELEEEGKPEGPANSSKTC 210
DB 175 GRRSGN-----GNKP--PSTDLKTC 192

RESULT 14
ADE58125 standard; protein; 203 AA.
XX
AC ADE58125;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein P51153, SEQ ID NO 3996.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002NO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
XX
PR 01-NOV-2001; 2001US-0346382P.
XX
.PR
```

```
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI: 2003-268312/26.
XX GENBANK; P51153.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 203 AA;

Query Match 48.1%; Score 532; DB 7; Length 203;
Best Local Similarity 47.6%; Pred. No. 8e-48;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDFPKMTIEVDGKVRIO 60
DB 1 MAKAYDHLFKLLIGDSGVGKTCLLIRFAEDNFNTYISTIGIDFKIRTVDIEGKKIKLQ 60

QY 61 IWDTAGOERYOTITKQYVRAAGIFLVYDISSERSYOHIMKWVSDVDEYAPEGVOKILIG 120
DB 61 VMDTAGOERFKTITTAIRYGAMGILVYDITDEKSFENIQNMWKSIVENASAGVERILLG 120

QY 121 NKADDEQKROVREGQOOLAKKEYGMDFYETSACTNLNIKESFTRITELVLQAHKKELEGL 180
DB 121 NKCDMEAKRKVKQEQADKLAREHGIRFFETSASKSMNVDEAFSSLDIL-----LKSG 174

QY 181 RMRASNELALAELEEEGKPEGPANSSKTC 210
DB 175 GRRSGN-----GNKP--PSTDLKTC 192

RESULT 15
AAB58196
ID AAB58196 standard; protein; 246 AA.
XX
XX AAB58196;
XX
XX 14-MAR-2001 (first entry)
XX
XX DT
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:05:27 ; Search time 23 Seconds  
(without alignments)  
475.857 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDLVRLLLIGDSVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCFUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	545	49.3	207	2	US-08-824-873-4
2	545	49.3	207	3	US-09-198-184-4
3	527.5	47.7	205	2	US-08-531-525-25
4	527.5	47.7	205	2	US-08-718-270A-25
5	518.5	46.9	198	2	US-08-531-525-51
6	518.5	46.9	198	2	US-08-718-270A-51
7	515.5	46.7	203	4	US-09-255-920A-12
8	503.5	45.6	207	2	US-08-531-525-35
9	503.5	45.6	207	2	US-08-718-270A-35
10	494.5	44.8	215	2	US-08-531-525-10
11	494.5	44.8	215	2	US-08-718-270A-10
12	488.5	44.2	194	2	US-08-531-525-34
13	488.5	44.2	194	2	US-08-718-270A-34
14	478.5	43.3	201	4	US-09-154-602-3
15	478.5	43.3	201	4	US-08-916-901-3
16	469	42.4	201	2	US-08-531-525-10
17	469	42.4	201	2	US-08-916-901-8
18	463	41.9	202	4	US-09-154-602-8
19	463	41.9	202	2	US-08-531-525-14
20	453.5	41.0	201	2	US-08-718-270A-14
21	453.5	41.0	201	2	US-08-531-525-13
22	452.5	41.0	190	2	US-08-718-270A-13
23	452.5	41.0	190	2	US-08-824-873-3
24	452.5	41.0	190	3	US-09-198-184-3
25	443	40.1	218	4	US-09-255-920A-7
26	443	40.1	218	2	US-08-531-525-19
27	430.5	39.0	217	4	US-08-718-270A-19
					US-09-255-920A-15

28	430	38.9	202	4	US-09-255-920A-11	Sequence 11, Appl
29	430	38.9	216	4	US-09-255-920A-14	Sequence 14, Appl
30	416	37.6	191	4	US-09-075-454-3	Sequence 3, Appl
31	412.5	37.3	212	4	US-09-399-913-67	Sequence 67, Appl
32	409.5	37.1	190	2	US-08-824-873-1	Sequence 1, Appl
33	409.5	37.1	190	3	US-09-198-184-1	Sequence 1, Appl
34	409.5	37.1	212	2	US-08-531-525-18	Sequence 18, Appl
35	409.5	37.1	212	2	US-08-718-270A-18	Sequence 18, Appl
36	409.5	37.1	214	2	US-08-531-525-52	Sequence 52, Appl
37	409.5	37.1	214	2	US-08-718-270A-52	Sequence 52, Appl
38	396.5	35.9	203	2	US-08-766-551-8	Sequence 8, Appl
39	396	35.8	192	4	US-09-255-920A-5	Sequence 5, Appl
40	388.5	35.2	213	4	US-09-976-594-1097	Sequence 1097, Ap
41	387	35.0	210	2	US-08-531-525-16	Sequence 16, Appl
42	387	35.0	210	2	US-08-718-270A-16	Sequence 16, Appl
43	373	33.8	213	2	US-08-773-423-8	Sequence 8, Appl
44	373	33.8	217	2	US-08-773-423-3	Sequence 3, Appl
45	372	33.7	208	2	US-08-531-525-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-08-824-873-4  
; Sequence 4, Application US/08824873  
; Patent No. 5843717  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,873  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 207 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 234746  
; US-08-824-873-4

Query Match 49.3%; Score 545; DB 2; Length 207;  
Best Local Similarity 52.9%; Pred. No. 2.4e-52;  
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDFKMKTIIEVDGKIRIQ 60  
Db 1 MAKYDYLFRLLIGDSGVGKTCVLFPSDFAFNSTFISGIDFKIRTIELDGKRIKQ 60  
QY 61 IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYQHIMKWSVDVDEYAPGQVKILIG 120  
Db 61 IWDTAGQERFTITATTAYRGAMGIMLVYDITNEKSPDNIRNWRINVEHASFADVERKMLG 120  
QY 121 NKADEQKQVGRGQOOLAKKEYGMDFYETSACTNINIKESFTRLTELVLQAHKKEGL 180  
Db 121 NKCDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKEGN 180  
QY 181 RMRASNE 187  
Db 181 SPQGSNQ 187

RESULT 2  
US-09-198-184-4  
; Sequence 4, Application US/09198184  
; Patent No. 6010859  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: NOVEL RAB PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/198,184  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/824,873  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0240 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 207 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 234746  
US-09-198-184-4

Query Match 49.3%; Score 545; DB 3; Length 207;  
Best Local Similarity 52.9%; Pred. No. 2.4e-52;  
Matches .99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;  
QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDFKMKTIIEVDGKIRIQ 60  
Db 1 MAKYDYLFRLLIGDSGVGKTCVLFPSDFAFNSTFISGIDFKIRTIELDGKRIKQ 60  
QY 61 IWDTAGQERYQTITKQYRRAQGIFLVYDISSERSYQHIMKWSVDVDEYAPGQVKILIG 120

Db 61 IWDTAGQERFTITATTAYRGAMGIMLVYDITNEKSPDNIRNWRINVEHASFADVERKMLG 120  
QY 121 NKADEQKQVGRGQOOLAKKEYGMDFYETSACTNINIKESFTRLTELVLQAHKKEGL 180  
Db 121 NKCDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFFTLARDIKAKMDKKEGN 180  
QY 181 RMRASNE 187  
Db 181 SPQGSNQ 187

RESULT 3  
US-08-531-525-25  
; Sequence 25, Application US/08531525  
; Patent No. 5840683  
; GENERAL INFORMATION:  
; APPLICANT: Hiavka, Joseph J.  
; APPLICANT: Pincus, Matthew R.  
; APPLICANT: No. 5840683le, John F.  
; APPLICANT: Abajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
; TITLE OF INVENTION: of P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/531,525  
; FILING DATE: 21-SEP-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feiber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 37-94  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 205 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Canis familiaris  
US-08-531-525-25

Query Match 47.7%; Score 527.5; DB 2; Length 205;  
Best Local Similarity 52.4%; Pred. No. 2.1e-50;  
Matches .97; Conservative 45; Mismatches 42; Indels 1; Gaps 1;  
QY 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDFKMKTIIEVDGKIRIQ 62  
Db 2 KYDYLFKLLIGDSGVGKTCVLFPSDFAFNSTFISGIDFKIRTIELDGKRIKQ 61  
QY 63 DTAGQERYQTITKQYRRAQGIFLVYDISSERSYQHIMKWSVDVDEYAPGQVKILGNK 122  
Db 62 DTAGQERFTITATTAY-RAMGIMLVYDITNEKSPDNIRNWRINVEHASFADVERKMLGNK 120  
QY 123 ADEQKQVGRGQOOLAKKEYGMDFYETSACTNINIKESFTRLTELVLQAHKKEGLRM 182

Db 121 CDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFPTLARDIKAKMDKKLEGNSP 180

Qy 183 RASNE 187

Db 181 QGSNQ 185

## RESULT 4

US-08-718-270A-25  
; Sequence 25, Application US/08718270A  
; Patent No. 5910478

GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5910478le, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptidomimetics Inhibiting  
TITLE OF INVENTION: the Oncogenic Action of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/718,270A  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/531,525  
FILING DATE: 21-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 33,878  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 78-95  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 205 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Canis familiaris

US-08-718-270A-25

Query Match 47.7%; Score 527.5; DB 2; Length 205;

Best Local Similarity 52.4%; Pred. No. 2.1e-50;

Matches 97; Conservative 45; Mismatches 42; Indels 1; Gaps 1;

Qy 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIENVGKIKVRIQ 62

Db 2 KTYDLFLKLLIGDSGVGKTCVLFPSDANSTFISITIGIDFKIRTTELQKRIKLQIW 61

Qy 63 DTAGQERYOTITKQYRRAQGIPLVYDISERSYOHIMKWSDVDYAPGQVKILGNK 122

Db 62 DTAGQERPTITTAY-RAMGIMLVYDITNEKSFNIRNWRINBEHASADVERKMLGNK 120

Qy 123 ADEQKQVGRQEQQLAKYGMDFYETSACTNINIKESFTRLRLVLAQHRKEGLRM 182

Db 121 CDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFPTLARDIKAKMDKKLEGNSP 180

Qy 183 RASNE 187

Db 181 QGSNQ 185

## RESULT 5

US-08-531-525-51  
; Sequence 51, Application US/08531525  
; Patent No. 5840683

GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 5840683le, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
TITLE OF INVENTION: of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Canis familiaris

US-08-531-525-51

Query Match 46.9%; Score 518.5; DB 2; Length 198;

Best Local Similarity 55.0%; Pred. No. 2e-49;

Matches 94; Conservative 45; Mismatches 31; Indels 1; Gaps 1;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIENVGKIKVRIQ 60

Db 1 MKKTYDLFLKLLIGDSGVGKTCVLFPSDDAFNTTIS-IGIDFKITVLEQKKIKLQ 59

Qy 61 IWDTAGQERYOTITKQYRRAQGIPLVYDISERSYOHIMKWSDVDYAPGQVKILIG 120

Db 60 IWDTAGQERPTITTYSYRGAMGIMLVYDITNGKSFENISKWLARNIDSHANEDVERMLLG 119

Qy 121 NKADEQKQVGRQEQQLAKYGMDFYETSACTNINIKESFTRLRLVLAQHRKEGLRM 171

Db 120 NKCDMDKVRVPKGEQIAREHGIRFETSAKVNNINIEKAFLTLAEDILR 170



RESULT 6  
US-08-718-270A-51  
; Sequence 51, Application US/08718270A  
; Patent No. 5910478  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.  
; APPLICANT: Pincus, Matthew R.  
; APPLICANT: No. 5910478le, John F.  
; APPLICANT: Abajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptidomimetics Inhibiting  
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/718,270A  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/531,525  
; FILING DATE: 21-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004,091  
; FILING DATE: 21-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ferber, Donna M.  
; REGISTRATION NUMBER: 33,878  
; REFERENCE/DOCKET NUMBER: 78-95  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Canis familiaris  
US-08-718-270A-51  
Query Match 46.9%; Score 518.5; DB 2; Length 198;  
Best Local Similarity 55.0%; Pred. No. 2e-49;  
Matches 94; Conservative 45; Mismatches 31; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRIQ 60  
Db 1 MKKTYDLLFKLLIGDSGVGKTCVLFPSDDAFNTTIS-IGIDFKIKTVLQKKIKIQ 59  
Qy 61 IWDTAGERYQITIKQYVRRAGQIFLVYDIISSERSYQHKMKVSDVDEYAPGKQKILIG 120  
Db 60 IWDTAGQERFHTITTSYVRGAMGLVYDITNGKSFENISKWLRNIDEHANEDVERMLLG 119  
Qy 121 NKADBEQKQVREGQQLAKEYGMDVYETSACTNLNLIKESFTRLTELVLQ 171  
Db 120 NKCDMDKRVVPGKGEGIAHGRFETSAKVNINIEKAFLLAEDILR 170  
RESULT 7

US-09-255-920A-12  
; Sequence 12, Application US/09255920A  
; Patent No. 6623980  
; GENERAL INFORMATION:  
; APPLICANT: Fisher, Joseph  
; APPLICANT: Lorens, James  
; APPLICANT: Anderson, David  
; APPLICANT: Luo, Ying  
; APPLICANT: Huang, Betty  
; APPLICANT: Shen, Mary  
; TITLE OF INVENTION: EXO1 and EXO2, EXOCYTOTIC PROTEINS  
; FILE REFERENCE: A65905-1/DJB/RMS  
; CURRENT APPLICATION NUMBER: US/09/255,920A  
; CURRENT FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: 60/075,534  
; PRIOR FILING DATE: 1998-02-23  
; PRIOR APPLICATION NUMBER: 60/086,650  
; PRIOR FILING DATE: 1998-05-26  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Mouse  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (34)  
; OTHER INFORMATION: The xaa at position 34 represents an unknown amino  
; OTHER INFORMATION: acid.  
US-09-255-920A-12  
Query Match 46.7%; Score 515.5; DB 4; Length 203;  
Best Local Similarity 56.0%; Pred. No. 4.4e-49;  
Matches 93; Conservative 41; Mismatches 31; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEF-HSHSHISTIGVDFKMTIEVDGKVR 59  
Db 1 MAKTYDYLFRLLIGDSGVGKTCVLFPSDDAFNSTFISTIGIDFKIRIELDKRIKL 60  
Qy 60 QIWDTAGERYQITIKQYVRRAGQIFLVYDIISSERSYQHKMKVSDVDEYAPGKQKILI 119  
Db 61 QIWDTAGQERFHTITTSYVRGAMGLVYDITNEKSFNIRNIRNEEHASADVERKML 120  
Qy 120 GNKADBEQKQVREGQQLAKEYGMDVYETSACTNLNLIKESFTRL 165  
Db 121 GNKCDVNDKQVSKERGEKLDYGIKFMETSAKANINVENAFTTL 166  
RESULT 8  
US-08-531-525-35  
; Sequence 35, Application US/08531525  
; Patent No. 5840683  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.  
; APPLICANT: Pincus, Matthew R.  
; APPLICANT: No. 5840683le, John F.  
; APPLICANT: Abajian, Henry B.  
; APPLICANT: Kende, Andrew S.  
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
; TITLE OF INVENTION: of P21 Ras  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 37-94  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORGANISM: Discopyge ommata  
US-08-531-525-35

Query Match 45.6%; Score 503.5; DB 2; Length 207;  
Best Local Similarity 46.4%; Pred. No. 9.8e-48;  
Matches 96; Conservative 56; Mismatches 44; Indels 11; Gaps 4;  
Qy 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDPKMKTIEVDGKVRQIW 62  
Db 2 KTYDYLFRLLIGDSGVGKTCCLFRFSEDAPNTFTISTIGDKIRIVELDGKIKLQIW 61  
Qy 63 DTAGQERYQTITKYYRRAQGIPLVYDISERSYOHIMKMWSDVDEVAPEGVQKILGNK 122  
Db 62 DTAGQERFRTIT-AAYRGAMGIMKV-DITNEKSFNKNWIRNIEEHASSDVERMILGNK 119  
Qy 123 ADEQKQVGRQEQQLAKYGMDFYETSACTNINIKESFRLTLVLQAHKLEGLRM 182  
Db 120 CDMNEKQVSKERGEKLAIDYGIKFLETSASKSINVEEAFITLARDIMTKLNKMM----- 174  
Qy 183 RASNELALAE--LEEEGKPEGPANSSK 208  
Db 175 ---NENSLQEAVDKLGKPPKPKSQKK 198

RESULT 9  
US-08-718-270A-35  
Sequence 35, Application US/08/18270A  
Patent No. 5910478  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 59104781e, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptidomimetics Inhibiting  
TITLE OF INVENTION: the Oncogenic Action of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/718,270A  
FILING DATE: 20-SEP-1996  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/531,525  
FILING DATE: 21-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,091  
FILING DATE: 21-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 78-95  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 207 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Discopyge ommata  
US-08-718-270A-35

Query Match 45.6%; Score 503.5; DB 2; Length 207;  
Best Local Similarity 46.4%; Pred. No. 9.8e-48;  
Matches 96; Conservative 56; Mismatches 44; Indels 11; Gaps 4;  
Qy 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDPKMKTIEVDGKVRQIW 62  
Db 2 KTYDYLFRLLIGDSGVGKTCCLFRFSEDAPNTFTISTIGDKIRIVELDGKIKLQIW 61  
Qy 63 DTAGQERYQTITKYYRRAQGIPLVYDISERSYOHIMKMWSDVDEVAPEGVQKILGNK 122  
Db 62 DTAGQERFRTIT-AAYRGAMGIMKV-DITNEKSFNKNWIRNIEEHASSDVERMILGNK 119  
Qy 123 ADEQKQVGRQEQQLAKYGMDFYETSACTNINIKESFRLTLVLQAHKLEGLRM 182  
Db 120 CDMNEKQVSKERGEKLAIDYGIKFLETSASKSINVEEAFITLARDIMTKLNKMM----- 174  
Qy 183 RASNELALAE--LEEEGKPEGPANSSK 208  
Db 175 ---NENSLQEAVDKLGKPPKPKSQKK 198

RESULT 10  
US-08-531-525-10  
Sequence 10, Application US/08531525  
Patent No. 5840683  
GENERAL INFORMATION:  
APPLICANT: Hlavka, Joseph J.  
APPLICANT: Pincus, Matthew R.  
APPLICANT: No. 58406831e, John F.  
APPLICANT: Abajian, Henry B.  
APPLICANT: Kende, Andrew S.  
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action  
TITLE OF INVENTION: of P21 Ras  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/531,525

```
;
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Arabidopsis thaliana
;
US-08-531-525-10

Query Match 44.8%; Score 494.5; DB 2; Length 215;
Best Local Similarity 46.9%; Pred. No. 1e-46;
Matches 100; Conservative 40; Mismatches 54; Indels 19; Gaps 4;

Qy 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVKRIQIWD 64
Db 12 YDYLKLLIGDSGVGKSCLLRSDGFTTFTTIGIDFKITIELDGKRIKLQIWD 71
Qy 65 AGQRYOTITKQYVRAQGIPLVYDVISERSYQHIMKWSVDVYAPGQVQKILGNKAD 124
Db 72 AQGR-RTITAYYRGAMGILLVYDVTDESFNNIRNWRNIEQHASDNVNVKILVGNKAD 130
Qy 125 -EEQKRVGREGQQQLAKEYGMDFYETSACTNINIKESFTLTLVLQAHKKELEGLMR 183
Db 131 MDESKRAVPTAKQALADEYGIKFFETSAKTNLNVVEVFFSIG-----RDIKQR 179
Qy 184 ASNELALAE-----LEEKGKPEGPANSSKTC 210
Db 180 LSDTDSRAEPATIKISQTDQAAGAQATQKSAC 212

RESULT 11
; Sequence 10, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 59104781e, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
```

```
;
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Arabidopsis thaliana
;
US-08-718-270A-10

Query Match 44.8%; Score 494.5; DB 2; Length 215;
Best Local Similarity 46.9%; Pred. No. 1e-46;
Matches 100; Conservative 40; Mismatches 54; Indels 19; Gaps 4;

Qy 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVKRIQIWD 64
Db 12 YDYLKLLIGDSGVGKSCLLRSDGFTTFTTIGIDFKITIELDGKRIKLQIWD 71
Qy 65 AGQRYOTITKQYVRAQGIPLVYDVISERSYQHIMKWSVDVYAPGQVQKILGNKAD 124
Db 72 AQGR-RTITAYYRGAMGILLVYDVTDESFNNIRNWRNIEQHASDNVNVKILVGNKAD 130
Qy 125 -EEQKRVGREGQQQLAKEYGMDFYETSACTNINIKESFTLTLVLQAHKKELEGLMR 183
Db 131 MDESKRAVPTAKQALADEYGIKFFETSAKTNLNVVEVFFSIG-----RDIKQR 179
Qy 184 ASNELALAE-----LEEKGKPEGPANSSKTC 210
Db 180 LSDTDSRAEPATIKISQTDQAAGAQATQKSAC 212

RESULT 12
; Sequence 34, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 58406831e, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; TITLE OF INVENTION: of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
```

```

; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Discopyge ommata
;
; US-08-718-270A-34
;
; Query Match 44.2%; Score 488.5; DB 2; Length 194;
; Best Local Similarity 55.8%; Pred. No. 4.1e-46;
; Matches 96; Conservative 41; Mismatches 30; Indels 5; Gaps 4
;
; QY 1 MAKQ-YDVLFRLLLIGDSGVGKTCLLCRFTDNEFHHSHISTIGVDFRMKKTIEVDGIKVR1 59
; DB 1 MAKKTYDLFLKLLIGDSGVGKTCVLFRRFSDDAFNTTFISTIGIDFKIKTVELHGKKIKL 60
;
; QY 60 QIWDTAQGRYQITTKQYHRAAGIFLVYDISERSYOHIMKWVSDVDVEYAPBGVQKILI 119
; DB 61 QIWDTAQGRFHFTIT-SYRGGANGIMLVYDITNAKSPENISKWLNRNIDEHANEDVERMLL 119
;
; QY 120 GNKADSEQRKQVREGOGQOLAKGYMGDFYTSACTNLNLIKESFTRLFELVLQ 171
; DB 120 GNK-DMEDKRVLLSKGQ--IAEHAIRFFETSAKANINIEKAFLTTLAEIDILQ 168

```

RESULT 14  
US-08-916-901-3  
Sequence 3, Application US/08916901  
Patent No. 5892012  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/916,901  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/POCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids

Sequence 3, Application US/08916901  
Patent No. 5892012  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Incyte Pharmaceuticals,  
STREET: 3174 Forter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fast-Seq for Windows Ver.ii  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/916,901  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids

CLONE: 2514506  
US-09-154-602-3  
Query Match 43.3%; Score 478.5; DB 4; Length 201;  
Best Local Similarity 45.0%; Pred. No. 5.5e-45;  
Matches 91; Conservative 42; Mismatches 64; Indels 5; Gaps 1;  
QY 1 MAKQYDVLFRLLLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDFKMKTIIEVDGKVRQ 60  
DB 1 MNPEYDYLFRLLLLIGDSGVGKSCLLLRFPADDTYTESYISTIGVDFKRTIIELDGKTIKQ 60  
QY 61 IWDTAGQERYQTITKQYRRAGIFLVYDISSERSYQHIMKMWSDVDVEYAPGVQKILIG 120  
DB 61 IWDTAGQERFRITTSYVRGAHGIIIVYDVTQESYANVKQWLOEIDRYASENVNKLVG 120  
QY 121 NKADEEQKRGVGRGQQLAKEYGMDFYETSACTNLNIKESFTLTELVLQAHKLEGL 180  
DB 121 NKSDLTKKVVNDTAKFADSLGIPFLETSAKNATNVEQAF-----MTMAAEIKKRMGP 175  
QY 181 RMRASNELALAELEEEBEGKPEG 202  
DB 176 GAASGGERPNLKIDSTPVKPKG 197

Search completed: March 15, 2004, 11:09:10  
Job time : 23 secs

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVRTUT04  
CLONE: 2514506  
US-08-916-901-3  
Query Match 43.3%; Score 478.5; DB 2; Length 201;  
Best Local Similarity 45.0%; Pred. No. 5.5e-45;  
Matches 91; Conservative 42; Mismatches 64; Indels 5; Gaps 1;  
QY 1 MAKQYDVLFRLLLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDFKMKTIIEVDGKVRQ 60  
DB 1 MNPEYDYLFRLLLLIGDSGVGKSCLLLRFPADDTYTESYISTIGVDFKRTIIELDGKTIKQ 60  
QY 61 IWDTAGQERYQTITKQYRRAGIFLVYDISSERSYQHIMKMWSDVDVEYAPGVQKILIG 120  
DB 61 IWDTAGQERFRITTSYVRGAHGIIIVYDVTQESYANVKQWLOEIDRYASENVNKLVG 120  
QY 121 NKADEEQKRGVGRGQQLAKEYGMDFYETSACTNLNIKESFTLTELVLQAHKLEGL 180  
DB 121 NKSDLTKKVVNDTAKFADSLGIPFLETSAKNATNVEQAF-----MTMAAEIKKRMGP 175  
QY 181 RMRASNELALAELEEEBEGKPEG 202  
DB 176 GAASGGERPNLKIDSTPVKPKG 197

RESULT 15  
US-09-154-602-3  
Sequence 3, Application US/09154602  
Patent No. 6300472  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: RAB PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,602  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/916,901  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0367 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 201 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVRTUT04

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:08:07 ; Search time 34 Seconds  
(without alignments)  
1316.603 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDLVFRLLIGDSGVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	212	9	US-09-817-198A-2
2	1105	100.0	401	9	US-09-764-868-701
3	1092	98.8	218	9	US-09-817-198A-5
4	1077	97.5	212	9	US-09-817-198A-4
5	832	75.3	188	9	US-09-764-868-1120
6	559	50.6	224	15	US-10-369-493-4997
7	540	48.9	221	15	US-10-264-049-2465
8	532	48.1	246	9	US-09-925-302-534
9	530.5	48.0	218	9	US-09-925-300-1571
10	528.5	47.8	207	9	US-09-794-257-8
11	523	47.3	206	14	US-10-128-714-3241
12	523	47.3	206	14	US-10-128-714-8600
13	514	46.5	201	14	US-10-179-766-6
14	514	46.5	201	15	US-10-369-493-5076
15	512	46.3	190	9	US-09-822-860-5

16	497	45.0	162	9	US-09-834-765-766	Sequence 766, Ap
17	478.5	43.3	201	9	US-09-967-736-3	Sequence 3, Appli
18	478.5	43.3	201	15	US-10-291-172-193	Sequence 193, App
19	478.5	43.3	224	14	US-10-102-806-466	Sequence 466, App
20	473.5	42.9	205	15	US-10-369-493-6261	Sequence 6261, Ap
21	473	42.8	210	14	US-10-032-585-7300	Sequence 7300, Ap
22	470.5	42.6	215	15	US-10-369-493-21876	Sequence 21876, A
23	469.5	42.5	225	14	US-10-128-714-8241	Sequence 8241, Ap
24	469	42.4	201	9	US-09-967-736-8	Sequence 8, Appli
25	467	42.3	222	9	US-09-820-003A-4	Sequence 4, Appli
26	462	41.8	207	14	US-10-032-585-7303	Sequence 7303, Ap
27	456.5	41.3	206	15	US-10-369-493-1787	Sequence 1787, Ap
28	448.5	40.6	216	15	US-10-369-493-5077	Sequence 5077, Ap
29	439	39.7	220	15	US-10-116-275-150	Sequence 150, App
30	434.5	39.3	198	9	US-09-794-257-16	Sequence 16, Appli
31	434.5	39.3	198	9	US-09-945-173-5	Sequence 5, Appli
32	434.5	39.3	198	9	US-09-972-529-4	Sequence 4, Appli
33	432	39.1	219	15	US-10-116-275-138	Sequence 138, App
34	430	38.9	223	9	US-09-817-199A-4	Sequence 4, Appli
35	429	38.8	223	9	US-09-817-199A-2	Sequence 2, Appli
36	429	38.8	223	15	US-10-108-260A-4746	Sequence 4746, Ap
37	427	38.6	226	9	US-09-764-868-684	Sequence 684, App
38	426	38.6	222	9	US-09-764-868-1106	Sequence 1106, Ap
39	421	38.1	214	15	US-10-369-493-5001	Sequence 5001, Ap
40	416	37.6	191	9	US-09-794-257-14	Sequence 14, Appli
41	416	37.6	191	13	US-10-051-986-3	Sequence 3, Appli
42	415.5	37.6	191	14	US-10-258-107-1	Sequence 1, Appli
43	413	37.4	223	15	US-10-369-493-1738	Sequence 1738, Ap
44	412.5	37.3	212	9	US-09-350-874-67	Sequence 67, Appli
45	412.5	37.3	212	14	US-10-106-989-67	Sequence 67, Appli

ALIGNMENTS

RESULT 1  
US-09-817-198A-2  
; Sequence 2, Application US/09817198A  
; Patent No. US20020146758A1  
; GENERAL INFORMATION:  
; APPLICANT: YE, Jane et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001188  
; CURRENT APPLICATION NUMBER: US/09/817,198A  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 212  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-817-198A-2

Query Match	100.0%;	Score 1105;	DB 9;	Length 212;
Best Local Similarity	100.0%;	Pred. No. 2.2e-106;		
Matches 212;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAKQYDLVFRLLIGDSGVGKTCLLCFTDNEFHSSHISTIGVDFKMTIIVDGIKVRIQ	60	
Db	1	MAKQYDLVFRLLIGDSGVGKTCLLCFTDNEFHSSHISTIGVDFKMTIIVDGIKVRIQ	60	
Qy	61	IWDTAGERYQTIKQYRRAQGIFLVYDIDISSERSYQHIMKWSVDVDAPEGVKILIG	120	
Db	61	IWDTAGERYQTIKQYRRAQGIFLVYDIDISSERSYQHIMKWSVDVDAPEGVKILIG	120	
Qy	121	NKADEEQRQVREGGQQLAKEYGMDFYETSACNLNLIKESFTRLTELVLQAHKKEGL	180	
Db	121	NKADEEQRQVREGGQQLAKEYGMDFYETSACNLNLIKESFTRLTELVLQAHKKEGL	180	
Qy	181	RMRASNELALAELEEEGKPEGPANSSKTCWC	212	
Db	181	RMRASNELALAELEEEGKPEGPANSSKTCWC	212	

```
Db      181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212

RESULT 2
US-09-764-868-701
; Sequence 701, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 701
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-701

Query Match      100.0%; Score 1105; DB 9; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.4e-106;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGIKVRIQ 60
Db      27 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGIKVRIQ 86

QY      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYVAPEGVKILIG 120
Db      87 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYVAPEGVKILIG 146

QY      121 NKADEEQKRVGREGQQQLAKYGMDFYETSACTNINIKESFTRLTTELVLQAHKKEGL 180
Db      147 NKADEEQKRVGREGQQQLAKYGMDFYETSACTNINIKESFTRLTTELVLQAHKKEGL 206

QY      181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
Db      207 RMRASNELALAELEEEGKPEGPANSSKTCWC 238

RESULT 3
US-09-817-198A-5
; Sequence 5, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Human
US-09-817-198A-5

Query Match      98.8%; Score 1092; DB 9; Length 218;
Best Local Similarity 97.2%; Pred. No. 5.2e-105;
Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGIKVRIQ 60
Db      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGIKVRIQ 60

QY      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYVAPEGVKILIG 120

Db      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYVAPEGVKILIG 120

Db      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYVAPEGVKILIG 120
QY      121 NKADEEQKRVGREGQQQLAKYGMDFYETSACTNINIKESFTRLTTELVLQAHKKEGL 180
Db      121 NKADEEQKRVGREGQQQLAKYGMDFYETSACTNINIKESFTRLTTELVLQAHKKEGL 180

QY      181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
Db      181 RTCASNELALAELEEDGKTEGPGANSSKTCWC 212

RESULT 4
US-09-817-198A-4
; Sequence 4, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-817-198A-4

Query Match      97.5%; Score 1077; DB 9; Length 212;
Best Local Similarity 97.6%; Pred. No. 1.8e-103;
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGIKVRIQ 60
Db      1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFPKMTIEVDGIKVRIQ 60

QY      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYVAPEGVKILIG 120
Db      61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDYVAPEGVKILIG 120

QY      121 NKADEEQKRVGREGQQQLAKYGMDFYETSACTNINIKESFTRLTTELVLQAHKKEGL 180
Db      121 NKADEEQKRVGREGQQQLAKYGMDFYETSACTNINIKESFTRLTTELVLQAHKKEGL 180

QY      181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
Db      181 RTCASNELALAELEEDGKTEGPGANSSKTCWC 212

RESULT 5
US-09-764-868-1120
; Sequence 1120, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (139)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```



NAME/KEY: SITE  
LOCATION: (151)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (161)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (164)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (188)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-1120

Query Match 75.3%; Score 832; DB 9; Length 188;  
Best Local Similarity 97.6%; Pred. No. 3.9e-78;  
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLIGDSVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQI 60  
Db 24 MAKQYDVLFRLLIGDSVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQI 83  
Qy 61 IWDTAGQERYQTITKQYRRAGGFLVYDIDISSERSYQHIMKWSVDVEYAPGQKILIG 120  
Db 84 IWDTAGQERYQTITKQYRRAGGFLVYDIDISSERSYQHIMKWSVDVEYAPGQKILIG 143  
Qy 121 NKADBEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTR 164  
Db 144 NKADBEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTR 187

RESULT 6  
US-10-369-493-4997  
Sequence 4997, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 4997  
LENGTH: 224  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-10-369-493-4997

Query Match 50.6%; Score 559; DB 15; Length 224;  
Best Local Similarity 55.3%; Pred. No. 1.1e-49;  
Matches 105; Conservative 42; Mismatches 39; Indels 4; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQI 60  
Db 1 MAKTYDYLKLLIGDSVGKTCVLFSEDAFNSTFTIGDKIRTIELDKRKLQI 60  
Qy 61 IWDTAGQERYQTITKQYRRAGGFLVYDIDISSERSYQHIMKWSVDVEYAPGQKILIG 120  
Db 61 IWDTAGQERYQTITKQYRRAGGFLVYDIDISSERSYQHIMKWSVDVEYAPGQKILIG 120  
Qy 121 NKADBEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHKLEGLR 180  
Db 121 NKCDIEERVRSDRGEQALIEYGTFLKTSKAKNLINIDEAFFTLARDI-----KSKMEQN 176  
Qy 181 RMASNEAL 190  
|||:|:

Db 177 EMRAATGAAI 186  
RESULT 7  
US-10-264-049-2465  
Sequence 2465, Application US/10264049  
Publication No. US20040005579A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA133P1  
CURRENT APPLICATION NUMBER: US/10/264,049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: Patent in Ver. 3.1  
SEQ ID NO 2465  
LENGTH: 221  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (2)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (6)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: MISC\_FEATURE  
LOCATION: (7)  
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-2465

Query Match 48.9%; Score 540; DB 15; Length 221;  
Best Local Similarity 52.7%; Pred. No. 9.7e-48;  
Matches 98; Conservative 45; Mismatches 43; Indels 0; Gaps 0;  
Qy 2 AKQYDVLFRLLIGDSVGKTCVLFSEDAFNSTFTIGDKIRTIELDKRKLQI 61  
Db 16 AKTYDYLKLLIGDSVGKTCVLFSEDAFNSTFTIGDKIRTIELDKRKLQI 75  
Qy 62 WDTAGQERYQTITKQYRRAGGFLVYDIDISSERSYQHIMKWSVDVEYAPGQKILIG 121  
Db 76 WDTAGQERYQTITKQYRRAGGFLVYDIDISSERSYQHIMKWSVDVEYAPGQKILIG 135  
Qy 122 KADBEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHKLEGLR 181  
Db 136 KCDVNDKRVSKERGERKALDYGIKFMETSAKININVENAFFTLARDIKAKMDKKLEGS 195  
Qy 182 MRASNE 187  
Db 196 PQGSNQ 201  
:|:

RESULT 8  
US-09-925-302-534  
Sequence 534, Application US/09925302  
Patent No. US20020044941A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA104  
CURRENT APPLICATION NUMBER: US/09/925,302  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05918  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 896





Db 1 MARRPYDMLFKLLIGDSGVGKTCILYRFSDDAFNTTFTSTIGIDPKIKTIELKGKIKL 60  
Qy 60 QIWDTAGQERYQTITKQYRRAGIFLVYDISSERSYOHIMKWVSDVDEYAPGVQKILI 119  
Db 61 QIWDTAGQERFHTTTSYRGAMGIMLVYDITNAKSPDNIAKWLNRNIDHASEDVVVMIL 120  
Qy 120 GNKADEEQKQVGRGQOQLAKYGMDFYTSACTNINIKESFTRLTELVLQHRKELEG 179  
Db 121 GNKCDMSDRRVSRERGEKIAQDHGISHFETSAKLNVHVDYAFYDLAEAIL----- 171  
Qy 180 LRWRASNELALAELEEEGKPEGPANSKTCWC 212  
Db 172 AKMPDSTD-----EQSRDTVNPVQPRQSSSGGC 200

## RESULT 15

US-09-822-860-5  
; Sequence 5, Application US/09822860  
; Patent No. US20020146795A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHU, Shiaoqing et al.  
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001214  
; CURRENT APPLICATION NUMBER: US/09/822.860  
; CURRENT FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Discopyge ommata  
US-09-822-860-5

Query Match 46.3%; Score 512; DB 9; Length 190;  
Best Local Similarity 47.7%; Pred. No. 6.3e-45;  
Matches 93; Conservative 52; Mismatches 42; Indels 8; Gaps 1;  
Qy 6 DVLFRLLIGDSGVGKTCILCRFTDNEFHSHISTIGVDFKMTIEVDGKVRQIWDTA 65  
Db 1 DYLKLLIGDSGVGKTCILFRSEDAFNTTFTSTIGIDPKIKTIELKGKIKIWDTA 60  
Qy 66 GQERYQTITKQYRRAGIFLVYDISSERSYOHIMKWVSDVDEYAPGVQKILIGNKADE 125  
Db 61 GQERFRTITAYYRGAMGIMKVDITNEKSPDNIAKWLNRNIDHASEDVVVMILGNKCDM 120  
Qy 126 EQKQVGRGQOQLAKYGMDFYTSACTNINIKESFTRLTELVLQHRKELEGLRMRAS 185  
Db 121 NEKQVSKERGEKLAIDYGIKFLKTSKSSINVEAFITLARDIMTKLNKKM----- 172  
Qy 186 NELALAELEEEGKP 200  
Db 173 NENSLQEAVDKIKSP 187

Search completed: March 15, 2004, 11:13:37  
Job time : 35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 11:03:18 ; Search time 2744 Seconds  
(without alignments)  
3348.660 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDLFRLLLTIGDSGVG.....LEEEKGKPGPANSKTCWC 212

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPRO\_spool\_p/US09817198/runat\_15032004\_101746\_20542/app\_query.fasta\_1.391  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817198 @CGN 1 1 5265 @runat\_15032004\_101746\_20542 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hcg.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_on.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_man.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1105	100.0	1054	6	AX399903 Sequence
2	1090	98.6	3139	10	BC027769 Mus muscu
3	1077	97.5	945	10	M83679 Sprague-Daw
4	1048.5	94.9	3326	9	BC040679 Homo sapi
5	938	84.9	2348	5	BC063736 Xenopus l
6	798	72.2	2560	10	BC013790 Mus muscu
7	797.5	72.2	2210	9	BX640825 Homo sapi
8	563.5	51.0	1726	5	BC053195 Danio rer
9	546.5	49.5	1613	3	AB006189 Drosophil
10	546.5	49.5	2349	3	AY060425 Drosophil
11	545	49.3	624	9	AF498943 Homo sapi
12	545	49.3	624	9	BT007184 Homo sapi
13	545	49.3	624	12	BT008275 Synthetic
14	545	49.3	660	9	X56741 H.sapiens m
15	545	49.3	760	4	X56385 Canine rab8
16	545	49.3	2048	9	BC002977 Homo sapi
17	545	49.3	2818	9	AK025165 Homo sapi
18	541	49.0	1337	10	BC019990 Mus muscu
19	539.5	48.8	765	5	M38391 Discopoge o
20	534.5	48.4	1211	9	BC000799 Homo sapi
21	534.5	48.4	1238	9	X75593 H.sapiens m
22	534.5	48.4	1356	5	BC060015 Xenopus l
23	532	48.1	612	9	AF498948 Homo sapi
24	531.5	48.1	840	5	M38390 Discopoge o
25	530.5	48.0	603	9	AF297660 Homo sapi
26	530.5	48.0	603	9	AF498945 Homo sapi
27	530.5	48.0	888	4	X56387 Canine rab1
28	530.5	48.0	897	10	AF035646 Mus muscu
29	530.5	48.0	1029	9	BC000896 Homo sapi
30	530.5	48.0	1071	9	AF086917 Homo sapi
31	530.5	48.0	3164	9	AF106681 Homo sapi
32	530.5	48.0	3185	10	BC056374 Mus muscu
33	530.5	48.0	3533	6	AX882713 Sequence
34	530.5	48.0	3533	6	BD159881 Primer fo
35	530.5	48.0	3533	9	AK023223 Homo sapi
36	530.5	48.0	3615	10	BC052735 Mus muscu
37	529.5	47.9	682	8	BT001952 Arabidops
38	529.5	47.9	708	8	BT002186 Arabidops
39	529.5	47.9	954	8	D01025 Arabidopsi
40	529.5	47.9	969	8	AY042795 Arabidops
41	529.5	47.9	1143	8	AY035132 Arabidops
42	528.5	47.8	624	6	AX236078 Sequence
43	528.5	47.8	740	10	RNU53475 Rattus norv
44	528.5	47.8	759	10	AF525280 Rattus no
45	528.5	47.8	1128	9	BC020654 Homo sapi

ALIGNMENTS

RESULT 1



508 AAGTGGGTCA GTGACGTGACGTGATGAGTACGCTCCAGAAGGAGTCCAGAAGATCCCTAATTGGG 567





US-09-817-198c-2 (1-212) x BC040679 (1-3326)	
Qy	1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyValGly 20
Db	5 ATGGCGAAGCAGTACGATGCTGTTCGGCTGTCTGCTGATCGGGGACTCCGGGTGGGC 64
Qy	21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
Db	65 AAGACCTGCCTGCTGTGCGCTTCACCGACAAACGAGTTCACCTCTCGCACATCTCCACC 124
Qy	41 IleglyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db	125 ATCCGTGTGATCTTAAGATGAACACCATAGAGTAGACGCCATCAAAAGTCGGATACAG 184
Qy	61 IleTpaAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db	185 ATCTGGACACATCGAGCGCAGAGATACCAACCATCACAAAGCAGTACTATCGCGGG 244
Qy	81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db	245 GCCCAGGGGATATTTTGGTCTATGACATTAGCAGCGAGCGCTCTTACCAGCACATCATG 304
Qy	101 LysTpaValSerAspValAspGlu----- 108
Db	305 AAGTGGGTCACTGACGTGGATGA- GGTAGGAGATGCCACCTCACTGCCGGGTGTGGAGA 363
Qy	108 ----- 108
Db	364 GGTGCTCACCAGGGAAGGCAAGCGGCGGAGAGAGCAATGCTTCAGGAA 423
Qy	109 -----TyrAlaProGluGlyValGlnLys 116
Db	424 GCTTTGCTCTCCACAGCCCTGGATGAAGACCTCTGTTAGCACCAGCAAGGCGTCCAGAAG 483
Qy	117 IleLeuIleGlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGly 136
Db	484 ATCTTTATTTGGGAATAAGCTGTATGAGGAGCAGAAACGCGAGGTGGGAAGAGACAAGGG 543
Qy	137 GlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeu 156
Db	544 CAGCAGCTGGCGAAGGAGTATGGCATGACTTCTATGAACAAGTGCCTGCACCAACCTC 603
Qy	157 AsnIleLysGluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGlu 176
Db	604 AACATTAAAGATCATTCACCGCTGTACAGAGCTGGTGTGTCGAGGCCCATAGGAGAG 663
Qy	177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
Db	664 CTGGAAGGCTCTCGGATGCGTGCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGAG 723
Qy	197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTpaCys 212
Db	724 GAGGGCAAAACCGAGGGCCCGAGCACTCTTCGAAAACCTCTCGTGTGC 771
RESULT 5	
LOCUS	BC063736 2348 bp mRNA linear VRT 09-DEC-2003
DEFINITION	Xenopus laevis cDNA clone MGC:68722 IMAGE:4057038, complete cds.
ACCESSION	BC063736
VERSION	BC063736.1 GI:39645074
KEYWORDS	MGC.
SOURCE	Xenopus laevis (African clawed frog)
ORGANISM	Xenopus laevis
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
AUTHORS	Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W. and Richardson, P.
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus Dev. Dyn. 225 (4), 384-391 (2002)
JOURNAL	
PUBMED	12454917

2 (bases 1 to 2348)

Klausner, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Strausberg, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Klausner, R.D., Collins, F.S., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Wallatton, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmeiz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

3 (bases 1 to 2348)

Klein, S. and Strausberg, R.

Direct Submission

Submitted (08-DEC-2003) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA

NIH-MGC Project

Contact: XGC help desk

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: Dr. Igor Dawid

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Gaithersburg Center (NISC), Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc\_mgc@nigr.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDevell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 128 Row: 1 Column: 12

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

Location/Qualifiers

1. 2348

/organism="Xenopus laevis"

/mol\_type="mRNA"

/db\_xref="taxon:8355"

/clone="MGC:68722 IMAGE:4057038"

/tissue\_type="Lung, adult Xenopus"

/clone\_lib="NICHG\_XGC\_Lul"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

138. .569

/codon\_start=1

/product="Unknown (protein for MGC:68722)"

/protein\_id="AAH63736.1"

/db\_xref="GI:39645075"

/translation="MAQYDVLFRLLIGDSVGKTCCLCRFTDNEFHPHSHTIGVD FKMTIEVDGIKVIQIWDTAQGRYQITKQYVYRAQGIPLVYDITSERYSQIMKW"

FEATURES

source

CDS

```

misc_feature      ASDVRYAPDGVQKILIGNKADBEQKRVQKNGKLFKCH"
159..551
/notes="RAB; Region: Rab subfamily of small GTPases"
/db_xref="CDD:cd00154"

ORIGIN
Alignment Scores:
Pred. No.:      1.07e-82      Length:      2348
Score:          938.00      Matches:    184
Percent Similarity: 85.28%      Conservative: 13
Best Local Similarity: 79.65%      Mismatches: 15
Query Match:     84.89%      Indels:    20
DB:              5          Gaps:      1

US-09-817-198c-2 (1-212) x BC063736 (1-2348)
QY      1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuLeuGlyAspSerGlyValGly 20
Db      138 ATGGCTAAACAATACGATGTTCTGTTCCGGCTGCTCTCATCGGGGACTCGGGGTCGGC 197
QY      21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeuSerThr 40
Db      198 AAGACCTGTTTGTGTGCAGATTCCCGACAATGAATTCACCCCTTCCACATCTCCACT 257
QY      41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyLysLeuValArgIleGln 60
Db      258 ATAGGAGTTGATTTCAAAATGAAGCAATAGAGTTGATGGATCAAGTACGAATACAA 317
QY      61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db      318 ATCTGGACACAGCTGCTCAAGAAGATATACAGACAATTAACCAACAGTAGTACTATAGAAGA 377
QY      81 AlaGlnGlyLeuPheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db      378 GCACAGGGTATATTCCTAGTCTATGATATATAACAGTAGTGAACGTTCTTACCAGCATATCATG 437
QY      101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db      438 AAATGGGCAGTGTATGATGATGATATGACCGGCGAGTTTCAGAAAATTTTAAATTTGGG 497
QY      121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGly----- 136
Db      498 AACAAAGCTGATGAAGAACAAACAAACACAGCTGTGGAAAAAATCAGGGGTTAAAGGCAATG 557
QY      137 -----GlnGlnLeuAlaLys 141
Db      558 TTCAACATTTGAACAGCATATTTGATTTGGACCATGAACATAAGGAGGAAAT-CTCGCAGAG 616
QY      142 GluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSer 161
Db      617 GAGTATGGGATGGACTTCTTTGAGACACAGTGCCTGCACATAATTACATATTAAGGAGTCT 676
QY      162 PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArg 181
Db      677 TTCACTCGTCTGACAGAGCTGCTCTGATGCTCACAAGAGGAGCTGGAGGGTTTCGA 736
QY      182 MetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysProGlu 201
Db      737 ATGCTCTCTGCTGATGAAGTGAATCTCGCTGAACCTGGAACGATGGCAGTAAATCAAA 796
QY      202 GlyProAlaAsnSerSerLysThrCysTrpCys 212
Db      797 GGAATGGAAAACTCTCTGAAATCTCTGTTGGTGC 829

RESULT 6
BC013790      2560 bp      mRNA      linear      ROD 11-DEC-2003
LOCUS      BC013790
DEFINITION      Mus musculus RAB15, member RAS oncogene family, mRNA (cDNA clone
MGC:6897 IMAGE:2655151), complete cds.
ACCESSION      BC013790
VERSION      BC013790.1 GI:15489393
KEYWORDS      MGC.
SOURCE      Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2560)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Scaletton,M., Soares,M.B., Donald,M.F., Casavant,T.L.,
Schepet,I.E., Brownstein,M.J., Urdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2560)
Strausberg,R.
Direct Submission
Submitted (04-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 5 Row: i Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19527265.
Location/Qualifiers
1. .2560
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="MGC:6897 IMAGE:2655151"
/tissue_type="Mammary tumor, MMTV-LTR/INT3 model. 5 month
old mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam2"
/lab_host="DH10E"
/notes="Vector: pCMV-SPORT6"
1. .2560
/genes="Rab15"
/db_xref="LocusID:104886"
/db_xref="MGI:1916865"
154..660
/codon_start=1
/product="Rab15 protein"
/protein_id="AAH13790.1"

gene
CDS
```

```

/db_xref="GI:15489394"
/db_xref="LocusID:104886"
translation="MAKQYDVLFRLLILGDSGVGKTCILCRFTDNFHHSHISTIGVD
FKMTIDVDGKIKVRIWDTAGQRYOTITKQYVRRAGQIFLVYDISSERSYQHMKW
VSDVDEYAPGVQKILIGNADEKQKRVGREGQQQLAKEYGMDFYETSACFFSPFS
ESTSPPL"
misc_feature
175..609
/Note="RAB; Region: Rab subfamily of small GTPases"
/db_xref="cd00154"

ORIGIN
Alignment Scores:
Pred. No.: 7,38e-69 Length: 2560
Score: 798.00 Matches: 152
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.35% Mismatches: 0
Query Match: 72.22% Indels: 0
DB: 10 Gaps: 0

US-09-817-198c-2 (1-212) x BC013790 (1-2560)
Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyValGly 20
Db 154 ATGGCGAAGAGTACAGTGTCTGTCGGCTACTGCTGATCGGGGACTCCGGGGTTGGC 213
Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db 214 AAGACATGCTGCTGTCGGCTTCCACGACACACGAGTTCCTCCTCGCATATCTCCACC 273
Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 274 ATCGGTGTTGACTTTAAGATGAAGATATCATGATAGACGGCATCAAAAGTGAAGAATACAG 333
Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 334 ATTGGGACACAGAGGCGGAGAGGTACAGACTATCAAAAGCAGTACTATCGGGGA 393
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 394 GCCCAGGGAATATTTTAGTCTAGACATTAGCATGAGCGCTCTCTATCAGCATATCATG 453
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 454 AAGTGGGTGTCAGTGCATGATGAGTACGCTCCAGAAAGAGTCCAGAAAGATCTTAATTGGG 513
Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140
Db 514 AATAAGGCTGATGAAGACGAAACGGCAGGTGGGGGAGAGCAGGGGCGCAGCTGGCT 573
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCys 153
Db 574 AAGAGTACGCGATGGACTTCTACGAACAAGTGCCTGC 612

RESULT 7
HSM806937 2210 bp mRNA linear PRI 28-AUG-2003
LOCUS HSM806937
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J06205 (from clone DKFZp686J06205).
ACCESSION BX640825
VERSION BX640825.1 GI:34365090
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2210)
AUTHORS Bloeker,H., Boecker,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
CONSRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (28-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686J06205) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://maps.gsfc.de/proj/cDNA/
Location/Qualifiers
1..2210
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="14q23.3"
/clone="DKFZp686J06205"
/tissue_type="human fetal brain"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="fetal"
1..2210
/gene="DKFZp686J06205"
<1..423
/gene="DKFZp686J06205"
/Note="similarity to GTP-binding protein rab15 (Rattus
norvegicus)"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE45901.1"
/db_xref="GI:34365091"
/translating="GPRPQGRNCQPLRLBAPEAGVRKSRGVDFPMKTIIEVDGIK
VRQIWDTAGQRYOTITKQYVRRAGQIFLVYDISSERSYQHMKWSDVDEGDATA
LPQCGGASPGKARRGPDGKANASRLCLPQPMWTKTSG"

ORIGIN
Alignment Scores:
Pred. No.: 6,9e-69 Length: 2210
Score: 797.50 Matches: 171
Percent Similarity: 50.89% Conservative: 0
Best Local Similarity: 50.89% Mismatches: 0
Query Match: 72.17% Indels: 165
DB: 9 Gaps: 1

US-09-817-198c-2 (1-212) x HSM806937 (1-2210)
Qy 42 GlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIle 61
Db 88 GGTGTTGACTTTAAGATGAAGACCATAGAGTAGACGCATCAAAAGTCGGATACAGATC 147
Qy 62 TrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAla 81
Db 148 TGGGACACTGCAGGCGAGGAGATACAGACCATCAAAAGCAGTACTATCGGGGCC 207
Qy 82 GlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLys 101
Db 208 CAGGGGATATTTTGGTCTATGACATTAGCAGCGAGCGCTCTTACCAGCATCATGAAG 267
Qy 102 TrpValSerAspValAsp
Db 268 TGGGTGAGTGCAGTGGATGAGGTAGGATGCCACCTCACTGCGCGGGGTGTGGAGAGGT 327
Qy 107
Db 328 GCCTCAGCGGGGAGGCAAGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 387
Qy 107
Db 388 TGCCTTCCACAGCCCTGGATGAAGACCTCTGGGTGAGTGAAGATGAGGAGGAGGAGGAG 447
Qy 107
Db 448 GCTGCCATGCCCTCACTCTCTTATACCTGCGAGCGCTCCACGGCTGTGTCTTTCCCGGAA 507
Qy 107

```

508 ATGAATTAGTTCACAGCTTTCCCTGTGAGCAGCTTCTTTCTGAAATCTTTGGGACCAGGT 567  
 107 ----- 107  
 568 GGAGTTCCAAGATTGGGATCTAGTCCCTGGCTCTGCACAATAGCTGTGGAGCCTTTGGGAAG 627  
 107 ----- 107  
 628 CCATTTGAATCCTCTGGGTCCCAGTTCTCTGTAGTAATGAGGGCTGGACTTTACATCCAATG 687  
 107 ----- 107  
 688 TCCTTTCCAGCTCTGATACCAAGTGGTCTAAACCAAGGAAGCACCAGTCTTTAGCCAGAGTG 747  
 108 ----- GluTyrAlaProGluGlyValGlnLys 116  
 748 TCTTCTACCTTAGCTCTCCCGGTGATACCTTTGAGTACGCACCAAGAGCGTCCAGAG 807  
 117 IleLeuIleGlyAsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGly 136  
 808 ATCTTATTGGGAATAAGGCTGTATGAGGAGCAGAGCGCAGGTGGGAGAGAGCAAGGG 867  
 137 GlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeu 156  
 868 CAGCAGCTGCGAGGAGTAGTGGCATGGACTTCTATGAAAACAAGTGGCTGCACCAACCTC 927  
 157 AsnIleLysGluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGlu 176  
 928 AACATTAAAGAGTCAATTCACGCGTCTCACAGAGCTGGTGTGTCAGGCCCATAGGAAGGAG 987  
 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGlu 196  
 988 CTGGAGGCCCTCCGATCGTGCAGCAATGAGTTGGCATCTGGCAGAGCTGGGAGGAG 1047  
 197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTyrCys 212  
 1048 GAGGGCAACCCAGGAGGCCAGCAACTCTCGAAAACTCTGGTGC 1095  
  
 RESULT 8  
 BC053195 1726 bp mRNA linear VRT 08-OCT-2003  
 LOCUS Danio rerio cDNA clone MGC:64010 IMAGE:6792909, complete cds.  
 DEFINITION  
 ACCESSION BC053195  
 VERSION BC053195.1 GI:31419200  
 KEYWORDS MGC  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 1726)  
 REFERENCE  
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultk, S.W.,  
 Viallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,  
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Touchman, J.W., Green, E.D.,  
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 TITLE  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```
Qy 61 IleTrrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 575 GTCTGGGACACAGCAGGACAGGAGGTTTAAGACCATCACCAGTGCATCTACAGAGGG 634
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 635 GCCATGGGAATCATCTCGTGTATGACATAATCTATGAGAAATCCTATGAGAAATTCAG 694
Qy 101 LysTrrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 695 AACTGGATGAAGACATCAAGAAATGATCAGCGGTGTAACTCGAATGTTACTGGGT 754
Qy 121 AsnLysAlaAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnLysAla 140
Db 755 AACAAAGTGTGACATTTAGGCGCAAGAGAAAGTCTCAAAAGAGAGCTGGTGAAGCTGCA 814
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 815 AAGAAACATGGTATTCGATTTCTGAGACTAGTGCCTCAATATGAGCATCAACGTTGAGGAG 874
Qy 161 SerPheThrArgLeuThr---GluLeuValLeuGlnAlaHisArgLys-----Glu 176
Db 875 TCTTTTACTTCTCTGTCGAGACATTTTGTGTAAGTCAATATGAGAGCCGGTCCATCT 934
Qy 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
Db 935 GGAGCTGAGGTCAAACTTTACCAGCAGCAGA-GAAAAAGTCTCTCTCCAAATGCTCTTCT 993
Qy 197 GluGlyLysPro 200
Db 994 TTAGGCAAAACCA 1005

RESULT 9
AB006189 1613 bp mRNA linear INV 07-AUG-1997
LOCUS Drosophila melanogaster mRNA for Rab10, complete cds.
DEFINITION AB006189
ACCESSION AB006189
VERSION AB006189.1 GI:2317271
KEYWORDS Rab10.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (sites)
SatoH.A.K., Tokunaga,F. and Ozaki,K.
Rab proteins of Drosophila melanogaster: novel members of the
Rab-protein family
FEBS Lett. 404 (1), 65-69 (1997)
JOURNAL 97228579
MEDLINE 9074639
PUBMED 9074639
REFERENCE 2 (bases 1 to 1613)
AUTHORS Ozaki,K.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-1997) Koichi Ozaki, Osaka University, Graduate
School of Science, Department of Biology; 1-1 Machikaneyama,
Toyonaka, Osaka 560, Japan (E-mail:ozaki@bio.sci.osaka-u.ac.jp,
Tel:+81-6-850-5439, Fax:+81-6-850-5439)
Location/Qualifiers
1. .1613
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Oregon R"
/db_xref="taxon:7227"
/tissue_type="head"
1. .1613
/gene="DRab10"
257. .871
/gene="DRab10"
/codon_start=1
/product="Rab10"
/protein_id="BAA21744.1"

FEATURES
source
```

```
/db_xref="GI:2317272"
/translaton="MAKTYDLPKLLIGDSGVGKTCILFRFSDDAFTSTFTSTIGI
DFKITYELRGKKIKLQWDTAQOERFHTTTSYRGAMGLMVDYITNEKSPFNIVK
WLRNIDSHANDEKEMILGNKCDMTDKRVNKRGEALAREHGRFMTSAKSNINIE
RAFCELAELDKTSRESAENQERVIIDRRNQEKAPGYSKCCA"

ORIGIN
Alignment Scores:
Pred. No.: 2,52e-44 Length: 1613
Score: 546.50 Matches: 105
Percent Similarity: 72.04% Conservative: 47
Best Local Similarity: 49.76% Mismatches: 50
Query Match: 49.46% Indels: 9
DB: 3 Gaps: 3
US-09-817-198C-2 (1-212) x AB006189 (1-1613)

Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyVal 19
Db 257 ATGCCAAGAAACCTACGATTTCCTCTTTAAACTGTGCTGATCGGTGATTCAGAGATG 316
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
Db 317 GGCAGAGAGTCATATTGTCGGTTCCTCGATGATGCATTCAGTCCACGTTCAATATCG 376
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIle 59
Db 377 ACCATAGCATCGATTTCAAATCAAAACAGTCGAGTCGCGCGCAAGAAGATCAAGCTG 436
Qy 60 GlnIleTrrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 79
Db 437 CAAATATGGGACACCGCGCGCAGGAGCGGTTCACACAGATAACCAACCATCTACTATCGA 496
Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
Db 497 GGCCTCATGGGCATATGCTGCTATGACATAACGAACAGAGAGAGTTTCGAGAACATA 556
Qy 100 MetLysTrrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
Db 557 GTCAATGGTTACGGAATATAGACGACGACGCGCAACGAGGATGTGGAGAAGATGATCCTC 616
Qy 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu 139
Db 617 GGCAACAAGTGGATATACGCGCAAGAGGGTGTCAACAGGAGCGCGCGCAAGCGATT 676
Qy 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db 677 GCCCGTGAACATGGCATTCGGTTCATGAAACATCCGCCAAGTCCGAACATAAATCATCGAG 736
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGly 179
Db 737 CGGCGCTTCTCGAGCTGGCGAGCGGCATCTCGAC-----AAGACATCAGGA 784
Qy 180 LeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLys 199
Db 785 CGCGAGTCGCGGAGAATCAGAGCGGTGATTCGATTCGCCGCGAACCAGGAGGAGCGC 844
Qy 200 ProGluGlyProAlaAsnSerSerLysThrCys 210
Db 845 CCG-----GGCTACAGCAAGTGTCTGC 865

RESULT 10
AB060425 2349 bp mRNA linear INV 05-NOV-2001
LOCUS Drosophila melanogaster LD39986 full length cDNA.
DEFINITION AB060425
ACCESSION AB060425
VERSION AB060425.1 GI:16648397
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
```











```

Db      190 ATATGGGACACAGCTGTCTAAGAACGGTTTCGGGACGATCACACAGCCTATTACAGGGGC 249
Qy      81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db     250 GCAATGGGCATCATGCTGGTCTATGACATCACCAACGAGAAATCCTTTGACAAATATATCCGG 309
Qy     101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db     310 AACTGGATTTCGAACATGAGGACATGCTTCTCCAGATGTCGAAAGATGATACTCGGA 369
Qy     121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGlnGlnGlyGlnGlnLeuAla 140
Db     370 AACAAAGTGTGATGTGAACGACAAAGACAAAGTTTCCAAAGGAACGGGAGAAAGCTGCC 429
Qy     141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db     430 CTGGACTATGGAATCAAGTTTCATGGAGACCAGTCCGAAGGCCAACATCAATGTGGAGAAC 489
Qy     161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
Db     490 GCATTTTCACTCTCGCCAGAGACATCAAGCAAAGATGGACAAAAAATTGGAAGGCCAAC 549
Qy     181 ArgMetArgAlaSerAsnGlu 187
Db     550 AGTCCCAAGGAGCAACCAG 570

```

Search completed: March 21, 2004, 12:57:23  
 Job time : 2753 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 11:01:53 ; Search time 354 Seconds  
(without alignments)  
2544.119 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAQYDVLFRLLIGDSGVG.....LEEEKGKPGANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09817198/runat\_15032004\_101745\_20514/app\_query.fasta\_1.391  
-DB=N Geneseq\_29Jan04 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817198 @CGN1\_1\_708 @runat\_15032004\_101745\_20514 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLDP=6 -DELEXT=7

Database :

- N\_Geneseq\_29Jan04.\*
- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002s.\*
- 7: Geneseqn2003as.\*
- 8: Geneseqn2003bs.\*
- 9: Geneseqn2003cs.\*
- 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	1054	6 ABQ93361	Abq93361 Human cDN
2	1105	100.0	2021	4 AAS27053	Aas27053 cDNA enco
3	1105	100.0	2021	9 ADB93231	Adb93231 Human cDN
4	1105	100.0	3257	7 AAD47168	Aad47168 Human Ras
5	1092	98.8	1666	3 AAC75813	Aac75813 Human ORF
6	832	75.3	566	4 AAS27472	Aas27472 cDNA enco
7	832	75.3	566	9 ADB93650	Adb93650 Human cDN
8	546.5	49.5	1540	4 ABL27707	AbL27707 Drosophil

9	545	49.3	1986	4	ABA09160	Aba09160 Human rab
10	540	48.9	2411	6	ABQ54410	Abq54410 Human ova
11	534.5	48.4	1274	3	AAF18072	Aaf18072 Lung canc
12	530.5	48.0	716	3	AAA40104	Aaa40104 Human Rab
13	530.5	48.0	861	3	AAA40108	Aaa40108 Human Rab
14	530.5	48.0	888	3	AAA40109	Aaa40109 Canine Ra
15	530.5	48.0	956	3	AAA96887	Aaa96887 Nucleotid
16	530.5	48.0	1537	3	AAF16196	Aaf16196 Human pro
17	530.5	48.0	3164	6	ABX04179	Abx04179 Human mRN
18	530.5	48.0	3533	4	AAH17889	Aah17889 Human CDN
19	529.5	47.9	674	5	AAS71453	Aas71453 DNA encod
20	528.5	47.8	1161	4	AAH75182	Aah75182 Nucleotid
21	528.5	47.8	1265	6	ABV78060	Abv78060 Hypoxia-r
22	528.5	47.8	2247	4	ABL29661	AbL29661 Drosophil
23	528.5	47.8	2497	4	AAS60878	Aas60878 Human can
24	528.5	47.8	2497	4	AAS60884	Aas60884 Human can
25	528.5	47.8	2497	4	AAS60893	Aas60893 Human can
26	528.5	47.8	2497	4	AAS60895	Aas60895 Human can
27	528.5	47.8	2497	5	ABV25781	Abv25781 Human pro
28	528.5	47.8	2497	5	ABV30037	Abv30037 Human pro
29	528.5	47.8	3077	4	AAH13912	Aah13912 Human CDN
30	523.5	47.4	775	9	ADC76178	Adc76178 DNA homol
31	523.5	47.4	775	9	ADD17377	AdD17377 DNA (SeqI
32	523	47.3	621	7	ABT21250	Abt21250 Aspergill
33	523	47.3	621	7	ABT19071	Abt19071 Aspergill
34	519.5	47.0	866	4	AAH04301	Aah04301 Human cDN
35	519	47.0	651	6	AB213031	Ab213031 Arabidops
36	519	47.0	1023	3	AAC51491	Aac51491 Arabidops
37	519	47.0	1023	3	AAC34347	Aac34347 Arabidops
38	513.5	46.5	1129	3	AAC44482	Aac44482 Zea mays
39	510.5	46.2	572	9	ADC76238	Adc76238 DNA homol
40	510.5	46.2	572	9	ADD17370	AdD17370 DNA (SeqI
41	510	46.2	911	3	AAC38429	Aac38429 Arabidops
42	509.5	46.1	709	9	ADD16498	AdD16498 DNA (SeqI
43	509	46.1	988	9	ADC75639	AdC75639 DNA homol
44	509	46.1	988	9	ADD16501	AdD16501 DNA (SeqI
45	507.5	45.9	720	9	ADD16526	AdD16526 DNA (SeqI

ALIGNMENTS

RESULT 1

ABQ93361

ID ABQ93361 standard; cDNA; 1054 BP.

XX AC ABQ93361;

XX AC ABQ93361;

XX DT 14-OCT-2002 (first entry)

XX XX

XX DE Human cDNA SEQ ID NO 74.

XX XX

KW Human; vulnary; dermatological; neuroprotective; nootropic; cancer;  
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;  
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;  
KW burn; central nervous system disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; immune disorder;  
KW autoimmune disorder; multiple sclerosis; diabetes; allergy; gene; ss.

XX OS Homo sapiens.

XX XX

XX WO200218424-A2.

XX XX

XX PD 07-MAR-2002.

XX XX

XX PF 31-AUG-2001; 2001WO-US027093.

XX XX

XX PR 01-SEP-2000; 2000US-00654935.

XX XX

XX PA (HYSE-) HYSEQ INC.

XX XX

XX PT Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;

XX PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;

XX XX



PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232402P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241211P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241825P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX P-PSDB; RAU17136.  
XX WPI: 2001-465460/50.  
XX  
PT Novel polypeptides useful for diagnosing, treating, preventing and/or  
PT prognosing disorders related to the proteins, including cancers, immune  
PT disorders and neuronal disorders.  
XX  
XX Claim 1; SEQ ID NO 88; 880pp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
XX diagnosing, preventing and treating diseases including immune system  
XX disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
XX disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
XX transplant rejections and graft versus host disease, infectious diseases  
XX (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
XX other blood-related disorders (sickle cell anaemia), myeloproliferative  
XX disorders, primary haematopoietic disorders, hyperproliferative disorders  
XX (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.  
XX Alzheimer's disease, Parkinson's disease), chromosomal abnormalities  
XX (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.  
XX glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),  
XX respiratory disorders, dermatological disorders, in wound healing,  
XX epithelial cell proliferation, endocrine disorders (e.g. Addison's  
XX disease), reproductive system disorders, gastrointestinal disorder  
XX (inflammatory disorders), liver disorders (cirrhosis), as stimulators of  
XX B-cell responsiveness to pathogens, activators of T-cells, to induce  
XX higher affinity antibodies, and as a means to induce tumour proliferation  
XX in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-  
XX AAS27850 represent novel signal transduction pathway protein coding  
XX sequences and PCR primers of the invention





CC acid encoding the polypeptide is useful for preventing, treating, or  
CC ameliorating a medical condition, which involves administering the  
CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid  
CC is useful for diagnosing a pathological condition or a susceptibility to  
CC a pathological condition in a subject, which involves determining the  
CC presence or absence of a mutation in the nucleic acid, and diagnosing a  
CC pathological condition or susceptibility to a pathological condition  
CC based on the presence or absence of the mutation. The polypeptide, the  
CC nucleic acid and an antibody to the polypeptide are useful for treating  
CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal  
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,  
CC the nucleic acid and the antibody are useful as immunosuppressive agents,  
CC as adjuvants to enhance immune responses, and as agents to induce higher  
CC affinity antibodies and increase serum immunoglobulin concentrations. The  
CC present sequence represents cDNA encoding a novel human protein. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format direct from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20020168711.

SQ Sequence 2021 BP; 464 A; 564 C; 553 G; 440 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 8.95e-118 Length: 2021  
Score: 1105.00 Matches: 212  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-817-198C-2 (1-212) x ADB93231 (1-2021)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly 20  
Db 81 ATGCGAAGCAGTACGATGCTGTTCCGGTGTCTGCTGATCGGGACTCGGGGTGGC 140  
Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeuSerThr 40  
Db 141 AAGACCTGCCTGCTGTGCGGTTCACCGACACGAGTTCACCTCTCGCACATCTCCACC 200  
Qy 41 IleGlyValAspPheLeuMetLysThrIleGluValAspGlyIleLysValArgIleGln 60  
Db 201 ATCCGTGTGTACTTAAGATGAACACCATAGAGGTAGCGGCATCAAGTCGGATACAG 260  
Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80  
Db 261 ATCTGGACACTGCAGGCGAGAGATACACAGACCATCACAGAGCAGTACTATCGGGC 320  
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100  
Db 321 GCCCAGGGGATATTTTGGTCTATGACATTAGCAGCGCGCTCTTACCGACACATCATG 380  
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120  
Db 381 AAGTGGGTGAGTGCAGTGGATGAGTACGCCACAGAGCGGCCAGAGATCTTATTGGG 440  
Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnLeuAla 140  
Db 441 AATAAGCTGATGAGGACCAACACGCGAGTGGAGAGAGAGAGGCGACAGCTGGCG 500  
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
Db 501 AAGGAGTATGTCATGGACTTCTATGAACAAGTGCCTGCACCACTCAACATTAAAGAG 560  
Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180  
Db 561 TCATTCAGCGCTGACAGACTGCTGTCAGGCCCATAGAGAGGAGCTGAGAGGCTC 620  
Qy 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysPro 200  
Db 621 CGGATGGTGCAGCAATGATTTGGCACTGCGACAGCTGGAGGAGGAGGAGGAGGAGGAG 680  
Qy 201 GluGlyProAlaAsnSerSerLysThrCysTrpCys 212

Db 681 GAGGGCCCGAGAACTCTTCGAAAACCTGCTGTGTC 716

RESULT 4

AAD47168

ID AAD47168 standard; cDNA; 3257 BP.

XX AAD47168;

AC AAD47168;

DT 24-FEB-2003 (first entry)

XX Human Ras-like protein encoding cDNA.

XX Human; Ras-like protein; inflammation; cell proliferation; apoptosis;  
XX immunodeficiency; neurodegenerative; Alzheimer's disease; cirrhosis;  
XX Parkinson's disease; wasting disease; cachexia; myocardial infarction;  
XX osteoporosis; atherosclerosis; glomerulonephritis; Crohn's disease;  
XX irritable bowel syndrome; multiple sclerosis; osteoarthritis; allergy;  
XX pancreatitis; autoimmune thyroiditis; rheumatoid arthritis; osteopathic;  
XX Sjogren's syndrome; infection; transgenic; gene therapy; nocrotropic; gout;  
XX neuroprotective; vasotropic; cytostatic; dermatological; nephrotropic;  
XX ophthalmological; tranquilliser; cancer; stroke; Grave's disease; AIDS;  
XX asthma; anaemia; drug screening; gene; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

5'UTR 1..44

FT /\*tag= a

FT CDS 45..683

FT /\*tag= b

FT 3'UTR 684..3257

FT /\*tag= c

FT WO200277193-A2.

XX 03-OCT-2002.

XX 27-MAR-2002; 2002WO-US009328.

XX 27-MAR-2001; 2001US-00817198.

XX (PEKE ) PE CORP.

XX Gan W, Ye J, Di Francesco V, Beasley EM;

XX WPI; 2003-018913/01.

XX P-PSDB; AAE29096.

New isolated human Ras-like protein polypeptide, useful for diagnosing,  
treating or preventing inflammation and disorders associated with cell  
proliferation and apoptosis, e.g. AIDS, cancer, allergies, anemia, asthma  
or stroke.

Claim 4; Page 73-74; 82pp; English.

The invention relates to human Ras-like protein and its corresponding  
nucleic acid. The Ras-like protein and DNA is useful in the development  
of human therapeutics and diagnostic compositions. They are useful in the  
diagnosis, prevention and treatment of inflammation and disorders  
associated with cell proliferation and apoptosis, e.g. AIDS and other  
infectious or genetic immunodeficiencies, neurodegenerative disease e.g.  
Alzheimer's disease, Parkinson's disease, wasting diseases e.g. cachexia,  
ischemic injuries e.g. myocardial infarction, stroke or reperfusion  
injury, toxin-induced diseases such as alcohol-induced liver damage or  
cirrhosis, osteoporosis or cancer. They are also used to treat disorders  
associated with inflammation including allergies, atopic dermatitis,  
atherosclerosis, asthma, anaemia, Crohn's disease, diabetes mellitus,  
Grave's disease, glomerulonephritis, gout, irritable bowel syndrome,  
lupus erythematosus, multiple sclerosis, osteoarthritis, pancreatitis,  
autoimmune thyroiditis, rheumatoid arthritis, Sjogren's syndrome,  
uveitis, trauma, or viral, bacterial, fungal, parasitic, protozoal or  
helminthic infections. The antibodies of the invention are useful in







```
RESULT 7
ID ADB93650 standard; cDNA; 566 BP.
XX
AC ADB93650;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA encoding a novel protein #497.
XX
KW ss; gene; human; autoimmune disease; Parkinson's disease; silicosis;
KW gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;
KW immunosuppressive agent; adjuvant; enhance immune response;
KW higher affinity antibody induction;
KW increased serum immunoglobulin concentration.
XX
OS Homo sapiens.
XX
PN US2002168711-A1.
XX
PD 14-NOV-2002.
XX
PF 17-JAN-2001; 2001US-00764868.
XX
31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0236868P.
PR 30-AUG-2000; 2000US-0238924P.
PR 01-SEP-2000; 2000US-0239287P.
PR 01-SEP-2000; 2000US-0239343P.
PR 01-SEP-2000; 2000US-0239344P.
PR 01-SEP-2000; 2000US-0239345P.
PR 05-SEP-2000; 2000US-0239509P.
PR 05-SEP-2000; 2000US-0239513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 27-SEP-2000; 2000US-0234997P.
PR 29-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
WPI; 2003-719985/68.
DR P-PSDB; ADB94263.
XX
PT New isolated polypeptide useful for diagnosing and treating
PT immunosuppressive conditions such as autoimmune disease and Parkinson's
PT disease.
XX
PS Claim 3; SEQ ID NO 507; 345pp; English.
XX
CC The invention relates to an isolated polypeptide. The polypeptide is
CC useful for diagnosing a pathological condition or a susceptibility to a
CC pathological condition in a subject, by determining the presence or
CC amount of expression of the polypeptide in a biological sample and
CC diagnosing a pathological condition or a susceptibility to a pathological
CC condition based on the presence or amount of expression of the
CC polypeptide. The polypeptide is also useful for identifying a binding
CC partner to the polypeptide, which involves contacting the polypeptide
CC with a binding partner and determining whether the binding partner
CC effects an activity of the polypeptide. The polypeptide or the nucleic
CC acid encoding the polypeptide is useful for preventing, treating, or
CC ameliorating a medical condition, which involves administering the
CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid
CC is useful for diagnosing a pathological condition or a susceptibility to
CC a pathological condition in a subject, which involves determining the
CC presence or absence of a mutation in the nucleic acid, and diagnosing a
CC pathological condition or susceptibility to a pathological condition
CC based on the presence or absence of the mutation. The polypeptide, the
CC nucleic acid and an antibody to the polypeptide are useful for treating
CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,
CC the nucleic acid and the antibody are useful as immunosuppressive agents,
CC as adjuvants to enhance immune responses, and as agents to induce higher
CC affinity antibodies and increase serum immunoglobulin concentrations. The
CC present sequence represents cDNA encoding a novel human protein. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format direct from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20020168711.
XX
SQ Sequence 566 BP; 140 A; 145 C; 172 G; 103 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 6.91e-87 Length: 566
Score: 832.00 Matches: 160
Percent Similarity: 97.56% Conservative: 0
Best Local Similarity: 97.56% Mismatches: 4
Query Match: 75.29% Indels: 0
DB: 9 Gaps: 0

US-09-817-198C-2 (1-212) x ADB93650 (1-566)
Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyValGly 20
Db 71 ATGGCGAAGCAGTACGATGTCTGCTCGGCTGTCTGTGATCGGGACTCCGGGTGGGC 130
Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db 131 AAGACCTGCCTGTGTGTCGGCTTCACCGACAAACGAGTTCCACTCTCTGCACATCTCCACC 190
Qy 41 IleGlyValAspPheLysMetLysThrIleGlyValAspGlyIleLysValArgIleGln 60
Db 191 ATCGGTGTGTGACTTTAAGATGAACCATAGAGGTAGACGCATCAAGTCCGGATACAG 250
Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
```

Db 251 ATCTGGGACACTGCGGGCAGGAGAGATACACACCATCAACAGCAGTACTATCGCGG 310  
Qy 81 AlaGlnGlyLeuPheLeuValTyArgPheLeuSerSerGluArgSerTyArgHisLeuMet 100  
Db 311 GCCCAGGGGATATTTTGGTCTATGACATTAGCAGCGCGCTCTTACCAGCACATCATG 370  
Qy 101 LysTrpValSerAspValAspGluTyAlaProGluGlyValGlnLysLeuLeuGly 120  
Db 371 AAGTGGGTCTAGTACGTGGATGATGACGACACAGAGCGCTCCAGAAATCTTATTGGG 430  
Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140  
Db 431 AATTAAGCTGATGACGAGCAGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 490  
Qy 141 LysGluTyArgMetAspPheTyGluThrSerAlaCysThrAsnLeuAsnLysLeuGlu 160  
Db 491 AANAGAGTATGGCATGGGACTTCTATGAAACAAAGTCCCTGCACCACTTAAACATTAAGAG 550  
Qy 161 SerPheThrArg 164  
Db 551 TCATTCACGCGT 562  
RESULT 8  
ABL27707  
ID ABL27707 standard; DNA; 1540 BP.  
AC ABL27707;  
XX  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 34594.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
XX 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
XX Claim 1; SEQ ID NO 34594; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1540 BP; 437 A; 353 C; 420 G; 330 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 3.3e-53 Length: 1540

Score: 546.50 Matches: 105  
Percent Similarity: 72.04% Conservative: 47  
Best Local Similarity: 49.76% Mismatches: 50  
Query Match: 49.46% Indels: 9  
DB: 4 Gaps: 3  
US-09-817-198c-2 (1-212) x ABL27707 (1-1540)  
Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuGlyAspSerGlyVal 19  
Db 257 ATGGCAAGAAACACCTACGATTGCTCTTTAACTGTTCTGATCGTTCATTCAGGAGTG 316  
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLysSer 39  
Db 317 GGCACAGCTGCATATTGTTCCGGTTCGCGATGATGCATTCACGTCCACGTTTCATATCG 376  
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgLys 59  
Db 377 ACCATAGGCATCGATTCAAAATCAAAACAGTCGAGCTGCGCGCAAGAGATCAAGCTG 436  
Qy 60 GlnIleTrpAspThrAlaGlyGlnGluArgTyArgGlnThrIleThrLysGlnTyArg 79  
Db 437 CAATATGGACACCCCGCGCAGAGCGGTTCCACAGATACCACTCGTACTATCGA 496  
Qy 80 ArgAlaGlnGlyLeuPheLeuValTyArgPheLysSerSerGluArgSerTyArgHisLys 99  
Db 497 GCGCCCATGGCATAATGCTGCTATGACATCAATCAACGAGAGAGAGTTCGAGAAACATA 556  
Qy 100 MetLysTrpValSerAspValAspGluTyAlaProGluGlyValGlnLysLeuLeu 119  
Db 557 GTCAAATGGTTACGGAATATAGACGAGCAGCGCAACGAGATGTGGAGAGATGATCCTC 616  
Qy 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeu 139  
Db 617 GGCACACAGTGCATATGACGACAGAGGGTGTCAACAGAGGCGCGCGAAGCGATT 676  
Qy 140 AlaLysGluTyArgGlyMetAspPheTyArgGluThrSerAlaCysThrAsnLeuAsnLys 159  
Db 677 GCGCGTGAACATGGCATTCGGTTTATGGAACATCCGCCCAAGTCGAAACATCAATCGAG 736  
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGly 179  
Db 737 CGGCGCTTCTGCGAGCTGCGCGGCGGCGGCTCTCTGAC-----AAGACATCAGGA 784  
Qy 180 LeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLys 199  
Db 785 CGCGAGTCGCGGAGAAATCAGAGCGCGTGTATTGATCGTCGCGGAACAGGAGAGCGG 844  
Qy 200 ProGluGlyProAlaAsnSerSerLysThrCys 210  
Db 845 CCG-----GGCTACAGCAAGTGTCTGC 865  
RESULT 9  
ABA09160  
ID ABA09160 standard; cDNA; 1986 BP.  
XX  
XX ABA09160;  
XX  
XX 11-JAN-2002 (first entry)  
XX  
XX Human rab8 homologue-encoding cDNA, SEQ ID NO:936.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; antiinflammatory;  
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;





```

PR 07-JUN-2000; 2000US-0209467P.
XX (HUMA-) HUMAN GENOME SCI. INC.
XX
XX Birse CE, Rosen CA;
XX WPI; 2002-147878/19.
DR P-PSDB; ABP41333.
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
XX
XX Claim 1; SEQ ID NO 290; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2411 BP; 600 A; 671 C; 583 G; 547 T; 0 U; 10 Other;
XX
Alignment Scores:
Pred. No.: 3.44e-52 Length: 2411
Score: 540.00 Matches: 98
Percent Similarity: 76.88% Conservative: 45
Best Local Similarity: 52.69% Mismatches: 43
Query Match: 48.87% Indels: 0
DB: 6 Gaps: 0

US-09-817-198c-2 (1-212) x ABQ54410 (1-2411)

Qy 2 AlalySglnTyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyValGlyIys 21
Db 47 GCGAAGACCTACGATTACCTGTTCAAGCTGCTGATCGGGGACTCGGGGGTGGGAAG 106
Qy 22 ThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisLeuSerThrIle 41
Db 107 ACCTGTGCTCTGTTCCGCTTCTCCGAGACGCCCTTCACTCCATTTATCTCCACCATA 166
Qy 42 GlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIle 61
Db 167 GGAATTGACTTTAAATTAAGACCATAGAGTCTGATGCGACAGAGAAATTAACCTCAGATA 226
Qy 62 TrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArgArgAla 81
Db 227 TGGGACACAGCGCGTCAAGAACCGGTTTCGGACGATCAACACGGGCTACTACAGGGGTGCA 286

```

```

Qy 82 GlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLys 101
Db 287 ATGGGCATCATGCTGCTTACGACATCACCAACGAGAAGTCTCTCGACAACATCCGGAAC 346
Qy 102 TrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsn 121
Db 347 TGGATTTCGCAACATTCAGGAGCAGCGCTCTGCAGACGTCGAAAAGATGATACTCGGGAAC 406
Qy 122 LysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLys 141
Db 407 AAGTGTGATGTGAATGACAGAGACAAAGTTTCCAGGAACGGGGAGAAAGCTTGGCCCTC 466
Qy 142 GluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSer 161
Db 467 GACTATGGAATCAAGTTTCATGGAGACCGCGGAGCCCAACATCAATGTGGAATAATGCA 526
Qy 162 PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGlyLeuArg 181
Db 527 TTTTTCACCTCTCCAGAGATATCAAAGCAAAATGACAAAATTTGGAAGGCAACAGC 586
Qy 182 MetArgAlaSerAsnGlu 187
Db 587 CCCAGGGGAGCAACACG 604
RESULT 11
AAF18072
ID AAF18072 standard; DNA; 1274 BP.
AC AAF18072;
XX 14-MAR-2001 (first entry)
XX
XX Lung cancer associated polynucleotide sequence SEQ ID 91.
XX
XX Human; lung cancer associated protein; neuroprotective; cytoskeletal;
XX cardioactive; immunomodulatory; muscular active; vulnerary;
XX gastrointestinal; nephrotropic; antiinfective; gynecological;
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX proliferative disorder; wound healing; infectious disease; ds.
XX
XX Homo sapiens.
XX
XX WC2000055180-A2.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005918.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Ruben SM;
XX
XX WPI; 2000-587514/55.
XX P-PSDB; AAB58196.
XX
XX Lung cancer associated gene sequences, referred to as lung cancer
XX antigens, useful for treatment, prevention, and diagnosis of disorders
XX such as lung cancer.
XX
XX Claim 1; Page 566; 1425pp; English.
XX
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
XX associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX associated proteins and polynucleotide sequences, their agonists, and
XX antagonists may have neuroprotective; cytoskeletal; cardioactive;
XX immunomodulatory; muscular active general; vulnerary; gastrointestinal
XX general; nephrotropic; antiinfective; gynecological; or antibacterial
XX activity. The invention also includes antibodies specific for the protein
XX or polynucleotide sequences. The lung cancer associated polynucleotide

```





```

DB: 3 Gaps: 1
US-09-817-198c-2 (1-212) x AAA40109 (1-888)
Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyVal 19
Db 133 ATGCGAAGAACGAGTACGACCTCTTTCAAGCTGCTCTCGATCGGAGACTCGGGAGTA 192
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
Db 193 GGGNAGACCTGGCGCTTTTTCGTTTCGATGATCCCTTCATACCACTTTATTTC 252
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIle 59
Db 253 ACCATAGGAATAGATTTTAAGATCAAAACAGTTCAATACAGGAAAGAGATCAAGCTA 312
Qy 60 GlnIleTrpAspThrAlaGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 79
Db 313 CAGATATGGGATACAGCAGGCCAGGCGGATTTTCACACCATCAACACCTCTCTACAGA 372
Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
Db 373 GGACCAATGGTATCATGCTAGTATATGACATCAACCAATGGTAAAGTTTGAACAATC 432
Qy 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
Db 433 AGCAATGGCTTAGAAACATAGATGAGCATGCCAATGAAGATGTGGAAGAATGTTACTA 492
Qy 120 GlyAsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu 139
Db 493 GGAACAAATGTGTATATGGACGATATAAAGAGTTGTCTACCTAAAGGAAAGGAGACGATT 552
Qy 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db 553 GCAAGGAGCATGTGTATTAGATTTTGTGAGACTAGTGCAAAAGTAATATAAATCAATCGAA 612
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db 613 AAGCCTTCTCTCACATTAGCTGAAGATATCTCTCGA 648

RESULT 15
AAA96887
ID AAA96887 standard; DNA; 956 BP.
XX
AC AAA96887;
XX
DT 19-FEB-2001 (first entry)
XX
DE Nucleotide sequence of human RAB10.
XX
KW RAB protein; GTPase; GTP binding; gene therapy; cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 127..729
FT /*tag= a
FT /product= "RAB10"
XX
XX WO200058464-A2.
XX
XX 05-OCT-2000.
XX
XX 13-MAR-2000; 2000WO-US006330.
XX
XX 25-MAR-1999; 99US-0126083P.
XX
XX (AXYS-) AXYS PHARM INC.
XX
XX Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;
XX WPI; 2000-647233/62.
XX P-PSDB; AAB19165.

```

```

XX
PT Novel isolated nucleic acid encoding a mammalian RAB protein useful for
PT identifying homologous or related genes, in producing composition that
PT modulates expression or function of RAB for cancer therapy.
XX
PS Claim 4; Page 38-39; 58pp; English.
XX
CC The present sequence encodes a mammalian RAB protein. RAB proteins
CC constitute the largest family of small GTPases, with over 40 identified
CC isoforms. RAB proteins contain four highly conserved peptide sequences
CC involved in GTP binding and hydrolysis. Compositions comprising RAB
CC nucleic acid are useful for identifying homologous or related genes, in
CC producing compositions that modulate the expression or function of RAB,
CC for gene therapy, mapping functional regions of the protein and in
CC studying associated physiological pathways. In addition, modulation of
CC the gene activity in vivo is used for prophylactic and therapeutic
CC purposes, such as treatment of cancer, and identification of cell type
CC based on expression. The DNA may also be used to identify expression of
CC the gene in a biological specimen
XX
SQ Sequence 956 BP; 294 A; 205 C; 213 G; 244 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.23e-51 Length: 956
Score: 530.50 Matches: 96
Percent Similarity: 82.56% Conservative: 46
Best Local Similarity: 55.81% Mismatches: 29
Query Match: 48.01% Indels: 1
DB: 3 Gaps: 1

US-09-817-198c-2 (1-212) x AAA96887 (1-956)
Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyVal 19
Db 127 ATGCGAAGAACGAGTACGACCTCTTTCAAGCTGCTCTCGATCGGAGACTCGGGAGTG 186
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
Db 187 GGGNAGACCTGGCGCTTTTTCGTTTCGATGATCCCTTCATACCACTTTATTTC 246
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIle 59
Db 247 ACCATAGGAATAGATTTTAAGATCAAAACAGTTCAATACAGGAAAGAGATCAAGCTA 306
Qy 60 GlnIleTrpAspThrAlaGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 79
Db 307 CAGATATGGGATACAGCAGGCCAGGCGGATTTTCACACCATCAACCTCTCTACAGA 366
Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
Db 367 GGGCAATGGTATCATGCTAGTATATGACATCAACCAATGGTAAAGTTTGAACAATC 426
Qy 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
Db 427 AGCAATGGCTTAGAAACATAGATGAGCATGCCAATGAAGATGTGGAAGAATGTTACTA 486
Qy 120 GlyAsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu 139
Db 487 GGAACAAATGTGTATATGGACGACAAAGAGTTGTCTACCTAAAGGAAAGGAGACAGATT 546
Qy 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db 547 GCAAGGAGCATGTGTATTAGATTTTGTGAGACTAGTGCAAAAGCAATATAAATCAATCGAA 606
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db 607 AAGCGTCTCTCACATTAGCTGAAGATATCTCTCGA 642

```

Search completed: March 21, 2004, 12:11:24

Job time : 362 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 04:59:32 ; Search time 83 Seconds  
(without alignments)  
1417.465 Million cell updates/sec

Title: US-09-817-198C-2  
Perfect score: 1105  
Sequence: 1 MAKQYDVLRLLLIGDSVG.....LEEBEGKEGPNSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlp  
-Q=/cgn2\_1/USPRO\_spool\_p/US09817198/runat\_15032004\_101743\_20457/app\_query.fasta\_1.391  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09817198.qcgn 1.1.85 @runat\_15032004\_101743\_20457 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	478.5	43.3	925	2	US-08-916-901-4
2	478.5	43.3	925	4	US-09-154-602-4
3	477.5	43.2	730	4	US-09-300-958A-16
4	467	42.3	723	4	US-09-016-434-1422
5	467	42.3	8137	4	US-09-566-921-7
6	456.5	41.3	1069	4	US-09-620-312D-646
7	437.5	39.6	1340	2	US-08-824-873-2
8	437.5	39.6	1340	3	US-09-198-184-2
9	429	38.8	875	4	US-09-075-454-10
10	429	38.8	1106	4	US-09-620-312D-959
11	429	38.8	2612	4	US-09-484-970B-142
12	424	38.4	833	4	US-09-620-312D-426

13	424	38.4	842	4	US-09-255-920A-6
14	419	37.9	639	4	US-09-399-913-66
15	411.5	37.2	970	3	US-08-888-077A-28
16	404.5	36.6	3936	4	US-09-919-172-49
17	388.5	35.2	1054	4	US-09-976-594-1096
18	374	33.8	1102	4	US-09-620-312D-828
19	373	33.8	847	2	US-08-773-423-4
20	366	33.1	944	4	US-09-016-434-1060
21	343.5	31.1	3745	4	US-09-776-594-387
22	343	31.0	890	3	US-08-741-411-4
23	340.5	30.8	803	4	US-09-075-454-13
24	339	30.7	912	4	US-09-016-434-1159
25	332.5	30.1	1984	4	US-09-023-655-7
26	321.5	29.1	1407	4	US-09-493-914-1
27	320	29.0	1172	4	US-09-075-454-8
28	320	29.0	1533	4	US-09-075-454-11
29	320	29.0	1886	4	US-09-620-312D-647
30	313	28.3	1255	2	US-08-766-551-6
31	297	26.9	606	4	US-09-016-434-1082
32	297	26.9	1175	2	US-08-773-423-6
33	296	26.8	921	4	US-09-016-434-1124
34	286.5	25.9	464	4	US-09-300-958A-14
35	283.5	25.7	624	4	US-09-016-434-1083
36	282.5	25.6	1084	4	US-09-673-395A-138
37	282	25.5	848	3	US-08-741-411-2
38	278.5	25.2	1393	4	US-09-391-741A-3
39	277	25.1	1631	4	US-09-620-312D-587
40	276.5	25.0	1393	4	US-09-391-741A-17
41	274.5	24.8	1749	4	US-09-149-476-54
42	273.5	24.8	1393	4	US-09-391-741A-27
43	272	24.6	1127	4	US-09-391-741A-1
44	271.5	24.6	820	3	US-08-741-411-6
45	271.5	24.6	820	4	US-09-016-434-112

ALIGNMENTS

RESULT 1  
US-08-916-901-4  
; Sequence 4, Application US/08916901  
; Patent No. 5892012  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: RAB PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/916,901  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0367 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166





Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140  
Db 426 AACAGAGCGACCTCACACCACCAAGAGTGTGGACCAACACACAGCAAGGAGTTTGCA 485  
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
Db 486 GACTCTCTGGGCATCCCTCTTCTGGAGACGAGCGCCCAAGATGCCCAATGTGCGACG 545  
Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlnLeu 180  
Db 546 GCGTTC-----ATGACCATGGCTGCTGAATCAAAAGCGGATGGGCGCT 590  
Qy 181 ArgMetArgAlaSerAsnGluLeuAlaGluLeuGluGluGluGluGluGluGluGlu 200  
Db 591 GGAGCAGCTCTGGGGGCGAGCGGCCCAATCTCAAGATCGACAGCACCCCTGTAAAGCG 650  
Qy 201 GluGly 202  
Db 651 GCTGGC 656  
RESULT 3  
US-09-300-958A-16  
; Sequence 16, Application US/09300958A  
; Patent No. 6495319  
; GENERAL INFORMATION:  
; APPLICANT: McClelland, Michael  
; APPLICANT: Welsh, John  
; APPLICANT: Trenkle, Thomas  
; TITLE OF INVENTION: Reduced Complexity Nucleic Acid Targets and Methods of  
; FILE REFERENCE: P-PH 3457  
; CURRENT FILING DATE: 1999-04-27  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR FILING DATE: 1998-08-27  
; PRIOR FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 730  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-300-958A-16  
Alignment Scores:  
Pred. No.: 6,36-52 Length: 730  
Score: 477.50 Matches: 87  
Percent Similarity: 72.34% Conservative: 49  
Best Local Similarity: 46.28% Mismatches: 51  
Query Match: 43.21% Indels: 1  
DB: 4 Gaps: 1  
US-09-817-198c-2 (1-212) x US-09-300-958A-16 (1-730)  
Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20  
Db 82 ATGCCCGGGGACTACGACCACTCTTCAAGTGTCTCATCGCGCAGCGGTGTGGGC 141  
Qy 21 LysThrCysLeuLeuCysAtgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40  
Db 142 AAGAGCAGTTTACTGTGGTGGTTTTCGAGACAACTTCTCAGGACGCTATCATCCACG 201  
Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60  
Db 202 ATCGAGTGGATTTCAGATCCGACCGGTGGAGATCAACGGGGAGAGGTGAACCTGACG 261  
Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80  
Db 262 ATCTGGGACACAGCGGGGCGAGGCGGCTTCGCGACCATCACCTCCAGTATTATCGGGG 321

Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100  
Db 322 ACCCAGCGGCTATTGTGGTTTACAGCTCACCAGTCCGAGTCTTTGTCAAGCTCAAG 381  
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120  
Db 382 CGGTGGCTTCACGAATCAACACCAAGACTGT---GATGATGTGTGCCGAATATTAGTGGT 438  
Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140  
Db 439 AATAAGATGACGACCTGAGCGGAGGTGTGGAGACGAAGATGCTACAAATTCGCC 498  
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
Db 499 GGGCAGATGGCATCCAGTTGTTCGAGCAGCGCCCAAGGAGATGTCAACGTGGAAGAG 558  
Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180  
Db 559 ATGTTCAACTGCATCAGGAGCTGCTCCGAGCAAAAGAAACAACCTGGCAAAACAG 618  
Qy 181 ArgMetArgAlaSerAsnGluLeu 188  
Db 619 CAGCAGCAACACAGACGATGTG 642  
RESULT 4  
US-09-016-434-1422  
; Sequence 1422, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1422:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 723 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: G550059  
US-09-016-434-1422  
Alignment Scores:







RESULT 10  
US-09-620  
; Sequenc

QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTyr 143  
Db 463 GATATGAGCAGCAAGAGTATCGCTTCGAAGACGAGACCTTGGCCAGGAGTAC 522  
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163  
Db 523 GGTGTTCCCTTCTCGAGACGAGCGCCCAAGACTGGCATGAATGTG----- 567  
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183  
Db 568 -----GAGTTAGCCTTCTTGCCCATCGCCCAAGGAA-----CTGAATACCGG 609  
QY 184 AlaSerAsnGlu 187  
Db 610 GCCGGGCATCAG 621

## RESULT 11

US-09-484-970B-142  
; Sequence 142, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkmut, Wayne  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; PRIORITY FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 142  
; LENGTH: 2612  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6426186 412477.1CB1  
US-09-484-970B-142

Alignment Scores:  
Pred. No.: 7,15e-45 Length: 2612  
Score: 429.00 Matches: 82  
Percent Similarity: 68.48% Conservative: 44  
Best Local Similarity: 44.57% Mismatches: 46  
Query Match: 38.82% Indels: 12  
DB: 4 Gaps: 3

## US-09-817-198C-2 (1-212) x US-09-484-970B-142 (1-2612)

QY 5 TyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24  
Db 91 TACGACCTCAGCGGAGGTGATGCTTCTGGAGACAGCGCTCGCAAAACATGTTTC 150  
QY 25 LeuCysArgPheThrAspAsnGluPheHisSer---SerHisIleSerThrIleGlyVal 43  
Db 151 CTGATCCAATTCAAGACGGGGCTTCTGTCCGGAACCTTCATAGCCACCGTCGGCATA 210  
QY 44 AspPheLysMetLysThrIleGluValAspGlyLysValArgIleGlnIleTyrAsp 63  
Db 211 GACTTCAGGAACAAGTGGTGAATGCTGATGCGGTGAGAGTGAAGCTGCAGATCTGGGAC 270  
QY 64 ThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGly 83  
Db 271 ACCGTGGGAGGACGGTTCGGAAGCGTCCACCATGCTTATTATACAGATGCTCAGGCC 330  
QY 84 IlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTyrVal 103  
Db 331 TTGCTTCTGCTGATGATGACATCAGCAACAATCTTCTTCGACAACATCAGGCGCTGGCTC 390  
QY 104 SerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAla 123  
Db 391 ACTGAGATTTCATGATGATGCCAGAGGACGCTGTGATCATGCTGCTAGGCAACAAGCGC 450

QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTyr 143  
Db 451 GATATGAGCAGCAAGAGTATCGCTTCGAAGACGAGACCTTGGCCAGGAGTAC 510  
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163  
Db 511 GGTGTTCCCTTCTCGAGACGAGCGCCCAAGACTGGCATGAATGTG----- 555  
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183  
Db 556 -----GAGTTAGCCTTCTTGCCCATCGCCCAAGGAA-----CTGAATACCGG 597  
QY 184 AlaSerAsnGlu 187  
Db 598 GCCGGGCATCAG 609

## RESULT 12

US-09-620-312D-426  
; Sequence 426, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chonghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunding  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghaast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIORITY FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIORITY FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIORITY FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 426  
; LENGTH: 833  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (95)..(745)  
US-09-620-312D-426

Alignment Scores:  
Pred. No.: 5.68e-45 Length: 833  
Score: 424.00 Matches: 83  
Percent Similarity: 60.98% Conservative: 42  
Best Local Similarity: 40.49% Mismatches: 68  
Query Match: 38.37% Indels: 12  
DB: 4 Gaps: 2

## US-09-817-198C-2 (1-212) x US-09-620-312D-426 (1-833)

QY 5 TyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24  
Db 101 TATGCTTATCTTCAAGTATATCATCATCGAGACAGCGTGGGGAAGTCAATGCTC 160  
QY 25 LeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThrIleGlyValAsp 44



Db	161	CTCTGTCAGTTTACAGATAAGGGTTCCAGCGCTCTCCAGACCTCAACATAGTGTGGAG	220
Qy	45	PhelysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIleTrpAspThr	64
Db	221	TTTGAGCTCGTATGGTCAACATTGATGGAAAAACAATCAACTGCAATCTGGGATACG	280
Qy	65	AlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlyIle	84
Db	281	GCTGGGCAAGAAATCCTCCGTTCTATCACCCGTTCTACTACAGGGGAGCAGCTGGAGCA	340
Qy	85	PhleuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSer	104
Db	341	CTGCTGGTGTAGCAATTACAGGCGTGAAACCTTCAACCACCTGACCTCATGGTTAGAG	400
Qy	105	AspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAsp	124
Db	401	GATGCCCGCAGCAGCTCTAGTTCCAAATGTTTATCATGCTCATTTGGGAATAGAGTGAC	460
Qy	125	GluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLysGluTyrGly	144
Db	461	CTAGAGTCCCGCAGCGATGTGAAGAGAGAAAGAGAGAGCGCTTGTCTAGGGAGCATGCA	520
Qy	145	MetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArg	164
Db	521	CTTATATTATCATGGAACCTTCAGCCAAAACAGCCCTGCAATGTTGAAGAGCGCTTCATTAA	580
Qy	165	LeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAla	184
Db	581	ACAGCCAAA-----GAATATATAGGAAGATCCAGCAGGGTTTA-----	619
Qy	185	SerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysProGluGlyProAla	204
Db	620	-----TTTGATGTCACAATGAGGCAAAATGGCATCAAGATTGGGCCCCCA	664
Qy	205	AsnSerSerLysThr	209
Db	665	CAGTCAATTTTCAACA	679
RESULT 13			
US-09-255-920A-6			
; Sequence 6, Application US/09255920A			
; Patent No. 6623980			
; GENERAL INFORMATION:			
; APPLICANT: Fisher, Joseph			
; APPLICANT: Lorens, James			
; APPLICANT: Anderson, David			
; APPLICANT: Luo, Ying			
; APPLICANT: Huang, Betty			
; APPLICANT: Shen, Mary			
; TITLE OF INVENTION: EXO1 and EXO2, EXOCYTIC PROTEINS			
; FILE REFERENCE: A65905-1/DJB/RMS			
; CURRENT APPLICATION NUMBER: US/09/255,920A			
; CURRENT FILING DATE: 1999-02-23			
; PRIOR APPLICATION NUMBER: 60/075,534			
; PRIOR FILING DATE: 1998-02-23			
; PRIOR APPLICATION NUMBER: 60/086,650			
; PRIOR FILING DATE: 1998-05-26			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: Patentin Ver. 2.0			
; SEQ ID NO 6			
; LENGTH: 842			
; TYPE: DNA			
; ORGANISM: Mouse			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: (3)			
; OTHER INFORMATION: The n at position 3 represents an unknown.			
; NAME/KEY: unsure			
; LOCATION: (6)..(8)			
; OTHER INFORMATION: The n at positions 6 through 8 represents an			
; OTHER INFORMATION: unknown.			
; NAME/KEY: unsure			





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 08:59:39 ; Search time 311 Seconds  
(without alignments)  
2521.609 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105  
Sequence: 1 MAXQYDVLFRLLIGSGVG.....LEBEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09817198/runat\_15032004\_101745\_20492/app\_query.fasta\_1.391  
-DB=Published Applications NA -Qfmt=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09817198 @CIGN 1.333 @runat\_15032004\_101745\_20492  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubnpa/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/ptodata/2/pubnpa/PCT\_NEW\_PUB.seq:
- 3: /cgn2\_6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq:
- 4: /cgn2\_6/ptodata/2/pubnpa/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/ptodata/2/pubnpa/US07\_NEW\_PUB.seq:
- 6: /cgn2\_6/ptodata/2/pubnpa/PCTUS\_PUBCOMB.seq:
- 7: /cgn2\_6/ptodata/2/pubnpa/US08\_NEW\_PUB.seq:
- 8: /cgn2\_6/ptodata/2/pubnpa/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/ptodata/2/pubnpa/US09A\_PUBCOMB.seq:
- 10: /cgn2\_6/ptodata/2/pubnpa/US09B\_PUBCOMB.seq:
- 11: /cgn2\_6/ptodata/2/pubnpa/US09C\_PUBCOMB.seq:
- 12: /cgn2\_6/ptodata/2/pubnpa/US09\_NEW\_PUB.seq:
- 13: /cgn2\_6/ptodata/2/pubnpa/US10A\_PUBCOMB.seq:
- 14: /cgn2\_6/ptodata/2/pubnpa/US10B\_PUBCOMB.seq:
- 15: /cgn2\_6/ptodata/2/pubnpa/US10C\_PUBCOMB.seq:
- 16: /cgn2\_6/ptodata/2/pubnpa/US10\_NEW\_PUB.seq:
- 17: /cgn2\_6/ptodata/2/pubnpa/US60\_NEW\_PUB.seq:
- 18: /cgn2\_6/ptodata/2/pubnpa/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
------------	-------	-------	-------	--------	-------	-------------

-----

1	1105	100.0	1054	12	US-10-363-616-74	Sequence 74, Appl
2	1105	100.0	2021	9	US-09-764-868-88	Sequence 88, Appl
3	1105	100.0	3257	9	US-09-817-198A-1	Sequence 1, Appl
4	832	75.3	566	9	US-09-764-868-507	Sequence 507, App
5	540	48.9	2411	15	US-10-264-049-290	Sequence 290, App
6	534.5	48.4	995	14	US-10-084-817-256	Sequence 256, App
7	534.5	48.4	1274	9	US-09-925-302-91	Sequence 91, Appl
8	530.5	48.0	1537	9	US-09-925-300-631	Sequence 631, Appl
9	530.5	48.0	3164	14	US-10-096-534-38	Sequence 38, Appl
10	528.5	47.8	624	9	US-09-794-257-9	Sequence 9, Appl
11	528.5	47.8	1161	9	US-09-794-257-7	Sequence 7, Appl
12	528.5	47.8	2497	9	US-09-834-975-879	Sequence 879, App
13	528.5	47.8	2497	9	US-09-834-975-885	Sequence 885, App
14	528.5	47.8	2497	9	US-09-834-975-896	Sequence 896, App
15	528.5	47.8	2497	9	US-09-834-975-896	Sequence 896, App
16	523	47.3	621	14	US-10-128-714-2241	Sequence 2241, Ap
17	523	47.3	621	14	US-10-128-714-7600	Sequence 7600, Ap
18	519	47.0	651	9	US-09-938-842A-836	Sequence 836, App
19	519	47.0	651	11	US-09-938-842A-836	Sequence 836, App
20	518	46.9	964	12	US-10-425-114-16115	Sequence 16115, A
21	514	46.5	1095	12	US-10-425-114-16115	Sequence 12044, A
22	513.5	46.5	1100	12	US-10-425-114-16320	Sequence 16320, A
23	513.5	46.5	1136	12	US-10-425-114-16320	Sequence 25934, A
24	513.5	46.5	1217	12	US-10-425-114-25934	Sequence 58714, A
25	513.5	46.5	1516	12	US-10-425-114-25934	Sequence 24388, A
26	512.5	46.4	880	12	US-10-425-114-24388	Sequence 31413, A
27	512.5	46.4	959	12	US-10-425-114-24388	Sequence 23752, A
28	512.5	46.4	1037	12	US-10-425-114-23752	Sequence 28854, A
29	512.5	46.4	1050	12	US-10-425-114-28854	Sequence 27513, A
30	511.5	46.3	1369	12	US-10-424-599-27513	Sequence 58679, A
31	507.5	45.9	1162	12	US-10-424-599-58679	Sequence 27209, A
32	506	45.8	1055	12	US-10-425-114-27209	Sequence 241, App
33	483	43.7	745	14	US-10-128-714-1241	Sequence 241, App
34	483	43.7	2745	14	US-10-128-714-241	Sequence 4, Appl
35	478.5	43.3	925	9	US-09-967-736-4	Sequence 48, Appl
36	478.5	43.3	939	14	US-10-102-806-48	Sequence 5, Appl
37	478.5	43.3	1898	15	US-10-221-278-5	Sequence 5, Appl
38	478.5	43.3	1898	15	US-10-291-172-5	Sequence 5, Appl
39	478.5	43.3	2078	14	US-10-119-428-5	Sequence 529, App
40	477.5	43.2	881	9	US-09-770-445-529	Sequence 43925, A
41	477	43.2	1144	12	US-10-424-599-43925	Sequence 20852, A
42	476.5	43.1	879	12	US-10-425-114-20852	Sequence 25791, A
43	476	43.1	766	12	US-10-425-114-25791	Sequence 2365, Ap
44	476	43.1	915	12	US-10-425-114-2965	Sequence 17495, A
45	476	43.1	943	12	US-10-425-114-17495	

ALIGNMENTS

RESULT 1  
US-10-363-616-74  
; Sequence 74, Application US/10363616  
; Publication No. US20040044181A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-113 (793)  
; CURRENT APPLICATION NUMBER: US/10363.616  
; CURRENT FILING DATE: 2003-03-03  
; PRIOR APPLICATION NUMBER: 09/654,935  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 490  
; SEQ ID NO 74  
; LENGTH: 1054  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (81)..(719)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1054)  
; OTHER INFORMATION: n = a,t,c or g









```
Db 247 ATCGGAATTGATTTCAAGATCCGCACTGTGGATATAGAGGGAAGAAGATCAAACTACAA 306
Qy 61 IletpAepThralaGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 307 GTCTGGGACACGGTGGCCAGAGCGGTTCAAGACATTAATCTACTGCTACTACCGTGA 366
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 367 GCCATGGGCATTATCTAGTATACCATCAGCATCCGGATGAGAAATCTTTCGAGAATATTCA 426
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 427 AACTGGGATGAAAGCATCAAGGAGAAATGCCCTCGGCTGGGTGGAGCGCTCTTCTGGGG 486
Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlyGlnLysAla 140
Db 487 AACAAATGTGACATGGAGCCCAAGAGGAGTGTGAGAGGAGCAGCGCGATAGTTGGCT 546
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 547 CGAGAGCATGGAATCCGATTTTCGAAACTAGTGTCTMAATCCAGTATGATGATGATGAG 606
Qy 161 SerPheThrArgLeuThrGluLeuValLeu----- 170
Db 607 GCCTTTAGTTCCTCGCCCGGACATCTTGCTCAAGTCAGAGGCCGAGATCAGGAAC 666
Qy 171 -----GlnAlaHisArgLysGlu 176
Db 667 GGCACAAAGCCTCCAGTACTGACCTGAAACTTGTGACAGAGAA-CACCAACAGTG 725
Qy 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
Db 726 CTCCTGGGCTGAGACCTTTCTTCCTCCCTCCCGGAGCTGAACCTGAGGGAGAC 785
Qy 197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211
Db 786 AACGGCAGAGGAGTGTAGCAGGGGAGAAATAGCAGAGGGGCTTGG 830

RESULT 7
US-09-925-302-91
; Sequence 91, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 91
; LENGTH: 1274
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1264)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1268)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-91

Alignment Scores:
Pred. No.: 4.1e-58 Length: 1274
Score: 534.50 Matches: 102
Percent Similarity: 66.38% Conservative: 54
Best Local Similarity: 43.40% Mismatches: 55
Query Match: 48.37% Indels: 25
```

```
DB: 9 Gaps: 1
US-09-817-198C-2 (1-212) x US-09-925-302-91 (1-1274)
Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
Db 132 ATGCCCAAGCCCTACGACCACCTCTTCAAGTGTCTGTGATCGGGGACTCGGGGGTGGGC 191
Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db 192 AAGACTTGTCTGATCATCTCGCTTTCAGAGGAGCAACTTCAACACACTTACATCTCCACC 251
Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 252 ATCGGAATTCATTTCAAGATCCGCACTGTGGATATAGAGGGAAGAAGATCAAACTACAA 311
Qy 61 IletpAepThralaGlnGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 312 GTCTGGGACACGGCTGGCCAGAGCGGTTCAAGACATTAATCTACTGCTACTACCGTGA 371
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 372 GCCATGGGCATTATCTAGTATACCATCAGCATCCGGATGAGAAATCTTTCGAGAATATTCA 431
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 432 AACTGGATGAAAGCATCAAGGAGATGCTCGGCTGGGTGGAGCGCTCTTCTGGGG 491
Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnLysAla 140
Db 492 AACAAATGTGACATGGAGGCCCAAGAGGAGTGTGAGAGGAGCGCGGATAGTTGGCT 551
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 552 CGAGAGCATGGAATCCGATTTTCGAAACTAGTGTCTMAATCCAGTATGATGATGATGAG 611
Qy 161 SerPheThrArgLeuThrGluLeuValLeu----- 170
Db 612 GCCTTTAGTTCCTCGCCCGGACATCTTGCTCAAGTCAGAGGCCGAGATCAGGAAC 671
Qy 171 -----GlnAlaHisArgLysGlu 176
Db 672 GGCACAAAGCCTCCAGTACTGACCTGAAACTTGTGACAGAGAA-CACCAACAGTG 730
Qy 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
Db 731 CTCCTGGGCTGAGACCTTTCTTCCTCCCGGAGCTGAACCTGAGGGAGAC 790
Qy 197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211
Db 791 AACGGCAGAGGAGTGTAGCAGGGGAGAAATAGCAGAGGGGCTTGG 835

RESULT 8
US-09-925-300-631
; Sequence 631, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 631
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
;
US-09-925-300-631
Alignment Scores:
Pred. No.: 1,75e-57 Length: 1537
Score: 530.50 Matches: 96
Percent Similarity: 82.56% Conservatives: 46
Best Local Similarity: 55.81% Mismatches: 29
Query Match: 48.01% Indels: 1
DB: 9 Gaps: 1

US-09-817-198C-2 (1-212) x US-09-925-300-631 (1-1537)
Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyVal 19
Db 533 ATGGCGAAGAAGACGTCAGACCTGCTTTTCAAGCTGCTCTGATCGGGGATTCGGGAGTG 592
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisLeSer 39
Db 593 GGGGAAGACCTCGCGCTTTTTCGGATGATGCTTCAATACCTTATTTTC 652
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgile 59
Db 653 ACCATAGGATAGACTTCAAGATCAAAACAGTTGAATTACAGGAAGAGATCAAGCTA 712
Qy 60 GlnIleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArg 79
Db 713 CAGATATGGGATACAGCAGCGCAGGAGCGATTTTCCACCATCATCAACCTCTCTACTACAGA 772
Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerGluArgSerTyrGlnHisIle 99
Db 773 GCGCAATGGGTATCATGCTAGTATATGACATCAACCAATGGTAAAGTTTGAACATC 832
Qy 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
Db 833 AGCAATGGCTTAGAACAATAGATGAGCATGCCAATGAGATGTGGAAAGAAATGTTACTA 892
Qy 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu 139
Db 893 GGAAACAAGTGTGATATGGACGACAAAGAGTTGTACCTTAAGGAAAGGAGAACAGATT 952
Qy 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db 953 GCAAGGGAGCATGGTATTTAGGTTTTCGAGACTAGTGCAAAAGCAAAATATAAACATCGAA 1012
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db 1013 AAGCGTTCCTCAGCTTAGCTGAGATATCTCTCGA 1048

RESULT 9
US-10-096-534-38
; Sequence 38, Application US/10096534
; Publication No. US20030166887A1
; GENERAL INFORMATION:
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Yates, Karen
; APPLICANT: Mizuno, Shuichi
; APPLICANT: Glowacki, Julie
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF SKELETAL DEGENERATION CONDITIONS
; FILE REFERENCE: B0801/7244/KA/ERP
; CURRENT APPLICATION NUMBER: US/10/096,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US 60/274,980
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-096-534-38
Alignment Scores:
Pred. No.: 5,02e-57 Length: 3164
Score: 530.50 Matches: 96
Percent Similarity: 82.56% Conservatives: 46
Best Local Similarity: 55.81% Mismatches: 29
Query Match: 48.01% Indels: 1
DB: 14 Gaps: 1

US-09-817-198C-2 (1-212) x US-10-096-534-38 (1-3164)
Qy 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyVal 19
Db 91 ATGGCGAAGAAGACGTCAGACCTGCTTTTCAAGCTGCTCTGATCGGGGATTCGGGAGTG 150
Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisLeSer 39
Db 151 GGGGAAGACCTCGCGCTTTTTCGGATGATGCTTCAATACCTTATTTTC 210
Qy 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgile 59
Db 211 ACCATAGGATAGACTTCAAGATCAAAACAGTTGAATTACAGGAAGAGATCAAGCTA 270
Qy 60 GlnIleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArg 79
Db 271 CAGATATGGGATACAGCAGCGCAGGAGCGATTTTCCACCATCATCAACCTCTCTACTACAGA 330
Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
Db 331 GCGCAATGGGTATCATGCTAGTATATGACATCAACCAATGGTAAAGTTTGAACATC 390
Qy 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
Db 391 AGCAATGGCTTAGAACAATAGATGAGCATGCCAATGAGATGTGGAAAGAAATGTTACTA 450
Qy 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu 139
Db 451 GGAAACAAGTGTGATATGGACGACAAAGAGTTGTACCTTAAGGAAAGGAGAACAGATT 510
Qy 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db 511 GCAAGGGAGCATGGTATTTAGGTTTTCGAGACTAGTGCAAAAGCAAAATATAAACATCGAA 570
Qy 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db 571 AAGCGTTCCTCAGCTTAGCTGAGATATCTCTCGA 606

RESULT 10
US-09-794-257-9
; Sequence 9, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
; FILE REFERENCE: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 624
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-794-257-9
Alignment Scores:
Pred. No.: 8,45e-58 Length: 624
Score: 528.50 Matches: 95
Percent Similarity: 80.68% Conservatives: 47
```







**THIS PAGE BLANK (USPTO)**



Result No.	Score	Query Match	Length	DB	ID	Description
1	1077	97.5	212	2	F42148	GTP-binding protein
2	559	50.6	224	2	T33855	hypothetical protein
3	545	49.3	207	2	B49647	GTP-binding protein
4	545	49.3	207	2	B36364	GTP-binding protein
5	532	48.1	203	2	A49647	GTP-binding protein
6	531.5	48.1	200	2	A38625	GTP-binding protein
7	530.5	48.0	200	2	D36364	GTP-binding protein
8	530	48.0	209	2	B38625	GTP-binding protein
9	526	47.6	206	2	I78851	GTP-binding protein
10	522.5	47.3	200	2	B42148	GTP-binding protein
11	519	47.0	216	2	T45901	GTP-binding protein
12	514	46.5	201	2	T28971	hypothetical protein
13	512	46.3	216	2	S33900	GTP-binding protein
14	512	46.3	216	2	T48378	GTP-binding protein
15	510	46.2	200	2	S12790	GTP-binding protein
16	509	46.1	215	2	T14565	GTP-binding protein
17	508.5	46.0	215	2	S57478	GTP-binding protein
18	508	46.0	216	2	J80640	GTP-binding protein
19	507.5	45.9	222	2	T14405	small GTP-binding
20	506	45.8	216	2	S57471	GTP-binding protein
21	505.5	45.7	217	2	S36365	GTP-binding protein
22	502.5	45.5	215	2	S57462	GTP-binding protein
23	501	45.3	203	2	S51495	GTP-binding protein
24	495.5	44.8	204	2	JC7589	Sec4p homolog - yeast
25	492.5	44.6	215	2	S57474	GTP-binding protein
26	483.5	43.8	208	2	A34716	GTP-binding protein
27	483	43.7	208	2	A38202	GTP-binding protein
28	482	43.6	203	2	S34253	GTP-binding protein
29	480	43.4	202	2	S38740	GTP-binding protein

A;Reference number: Z21424  
A;Accession: T33855  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-224 <LEAD>  
A;Cross-references: EMBL:AF106592; PIDN:AC78494.1; GSPDB:GN00019; CESP:D1037.4  
A;Experimental source: strain Bristol N2; clone D1037.  
C;Genetics:  
A;Gene: CESP:D1037.4  
A;Map position: 1  
A;Introns: 10/3; 62/2; 82/3; 181/1  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 50.6%; Score 559; DB 2; Length 224;  
Best Local Similarity 55.3%; Pred. No. 2.4e-37;  
Matches 105; Conservative 42; Mismatches 39; Indels 4; Gaps 1;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKIRIQ 60  
Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDDSFNPSITIGIDFKIRTIELDGKIKLQ 60

Qy 61 IWDTAGOERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWSDVDVEYAPGQKILIG 120  
Db 61 IWDTAGOERFTITTYIRGANGIMLVYDITNEKSFNIRNWIIEHASADVERKILG 120

Qy 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Db 121 NKCDIEERVRSDRGQLAIEYGTKEFLETSAKANLIDEAFFTLARDI---KSKMEQN 176

Qy 181 RMRASNEAL 190  
Db 177 EMRAATGAAI 186

## RESULT 3

GTP-binding protein rab8 - human  
C;Species: Homo sapiens (man)  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Jan-2001  
C;Accession: B49647; S36817  
R;Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavittian, A.; Louvar  
J. Cell Biol. 124, 101-115, 1994  
A;Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cell  
A;Reference number: A49647; MUID:94124602; PMID:8294494  
A;Accession: B49647  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-207 <ZAH>  
A;Cross-references: EMBL:X56741; NID:g452317; PIDN:CAA40065.1; PID:g452318  
R;Joberty, G.; Tavittian, A.; Zahraoui, A.  
FEBS Lett. 330, 323-328, 1993  
A;Title: Isoprenylation of Rab proteins possessing a C-terminal CaaX motif.  
A;Reference number: S36817; MUID:93387463; PMID:8375503  
A;Accession: S36817  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 175-186 <JOB>  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;9-124/Domain: translation elongation factor Tu homology <ETU>  
F;15-22/Region: nucleotide-binding motif A (P-loop)  
F;121-124/Region: GTP-binding NKXD motif  
F;151-153/Region: GTP-binding SAK/L motif  
F;204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 49.3%; Score 545; DB 2; Length 207;  
Best Local Similarity 52.9%; Pred. No. 2.8e-36;  
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKIRIQ 60  
Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDDSFNPSITIGIDFKIRTIELDGKIKLQ 60

Qy 61 IWDTAGOERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWSDVDVEYAPGQKILIG 120  
Db 61 IWDTAGOERFTITTYIRGANGIMLVYDITNEKSFNIRNWIIEHASADVERKILG 120  
Qy 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Db 121 NKCDVNDKRVSKERGEKLDYGIKFMETSAKANINVENAFTTLARDIKAKMDKKLEGN 180  
Qy 181 RMRASNE 187  
Db 181 SPOGSNQ 187

## RESULT 4

B36364  
GTP-binding protein rab8 - dog  
C;Species: Canis lupus familiaris (dog)  
C;Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 02-Feb-2001  
C;Accession: B36364; S15604  
R;Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.  
Mol. Cell. Biol. 10, 6578-6585, 1990  
A;Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.  
A;Reference number: A36364; MUID:91061765; PMID:2123294  
A;Accession: B36364  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-207 <CHA>  
A;Cross-references: GB:X56385; NID:g920; PIDN:CAB56776.1; PID:g6006436  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;9-124/Domain: translation elongation factor Tu homology <ETU>  
F;15-22/Region: nucleotide-binding motif A (P-loop)  
F;121-124/Region: GTP-binding NKXD motif  
F;151-153/Region: GTP-binding SAK/L motif  
F;204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 49.3%; Score 545; DB 2; Length 207;  
Best Local Similarity 52.9%; Pred. No. 2.8e-36;  
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKIRIQ 60  
Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDDSFNPSITIGIDFKIRTIELDGKIKLQ 60

Qy 61 IWDTAGOERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWSDVDVEYAPGQKILIG 120  
Db 61 IWDTAGOERFTITTYIRGANGIMLVYDITNEKSFNIRNWIIEHASADVERKILG 120

Qy 121 NKADDEQKRVGREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEGL 180  
Db 121 NKCDVNDKRVSKERGEKLDYGIKFMETSAKANINVENAFTTLARDIKAKMDKKLEGN 180

Qy 181 RMRASNE 187

Db 181 SPOGSNQ 187

## RESULT 5

A49647  
GTP-binding protein Rabi3 - human  
C;Species: Homo sapiens (man)  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 02-Feb-2001  
C;Accession: A49647  
R;Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavittian, A.; Louvar  
J. Cell Biol. 124, 101-115, 1994  
A;Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cell  
A;Reference number: A49647; MUID:94124602; PMID:8294494  
A;Accession: A49647  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-203 <ZAH>  
A;Cross-references: EMBL:X75593; NID:g452319; PIDN:CAA53266.1; PID:g452320  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;9-124/Domain: translation elongation factor Tu homology <ETU>  
F;15-22/Region: nucleotide-binding motif A (P-loop)  
F;121-124/Region: GTP-binding NKXD motif  
F;151-153/Region: GTP-binding NKXD motif  
F;212-22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta  
F;200/Binding site: farnesyl (Cys) (covalent) #status predicted

Query Match 48.1%; Score 532; DB 2; Length 203;  
Best Local Similarity 47.6%; Pred. No. 3e-35;  
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDYDFKMKTIIEVDGKIVRIQ 60  
Db 1 MAKAYDHLFKLLIGDSGVGKTCIIIRPAEDFNNTVISTIGIDFKIRTVDIEGKKIKLQ 60  
Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWVSDVDVEYAPGKILIG 120  
Db 61 VMDTAGQERFHTITAYYRGAMGILVYDITDEKSFENIQNMWKSIIKENASAGVERLLLG 120  
Qy 121 NKADDEOKRVGREGQOOLAKYGMDFYETSACTNLNIKESFTRLTELVLQAHKKELEGL 180  
Db 121 NKCDMEARKVKQKQADKLAREHGIRFFETSAKSMNVDEAFSLARDIL-----LKSG 174  
Qy 181 RMRASNELALAELEEKGKPGPANSKTC 210  
Db 175 GRRSGN-----GNKP--PSTDLEKTC 192

## RESULT 6

A38625  
GTP-binding protein oral - electric ray (Discopyge ommata)  
C;Species: Discopyge ommata  
C;Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 02-Feb-2001  
C;Accession: A38625  
R;Ngsee, J.K.; Elferink, L.A.; Scheller, R.H.  
J. Biol. Chem. 266, 2675-2680, 1991  
A;Title: A family of ras-like GTP-binding proteins expressed in electromotor neurons.  
A;Reference number: A38625; MUID:91115900; PMID:1899244  
A;Accession: A38625  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-200 <NGS>  
A;Cross-references: GB:M38390; NID:g213114; PIDN:AAA49230.1; PID:g213115  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; membrane trafficking; nucleotide binding; P-loop  
F;10-125/Domain: translation elongation factor Tu homology <ETU>  
F;16-23/Region: nucleotide-binding motif A (P-loop)  
F;122-125/Region: GTP-binding NKXD motif  
F;152-154/Region: GTP-binding SAK/L motif

Query Match 48.1%; Score 531.5; DB 2; Length 200;  
Best Local Similarity 56.4%; Pred. No. 3.2e-35;  
Matches 97; Conservative 44; Mismatches 30; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDYDFKMKTIIEVDGKIVRI 59  
Db 1 MAKTYDHLFKLLIGDSGVGKTCVLFPSDAFNTTISTIGIDFKIKTVLHGKKIKL 60  
Qy 60 QIWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWVSDVDVEYAPGKILIG 119  
Db 61 QIWDTAGQERFHTITAYYRGAMGILVYDITNAKSFENISKWLNRNDEHANEDVERMLL 120  
Qy 120 GNKADEOKRVGREGQOOLAKYGMDFYETSACTNLNIKESFTRLTELVLQ 171  
Db 121 GNKCMEDKRVLSKGRQIAREHAIRFFETSAKANINIEKAFTLTAEIDILQ 172

## RESULT 7

D36364  
GTP-binding protein rab10 - dog  
C;Species: Canis lupus familiaris (dog)  
C;Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 02-Feb-2001

C;Accession: D36364; S15601  
R;Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.  
Mol. Cell. Biol. 10, 6578-6585, 1990  
A;Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.  
A;Reference number: A36364; MUID:91061765; PMID:2123294  
A;Accession: D36364  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-200 <CHA>  
A;Cross-references: GB:X56387; NID:g914; PIDN:CAA39798.1; PID:g915  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;10-125/Domain: translation elongation factor Tu homology <ETU>  
F;16-23/Region: nucleotide-binding motif A (P-loop)  
F;122-125/Region: GTP-binding NKXD motif  
F;152-154/Region: GTP-binding SAK/L motif  
F;199,200/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 48.0%; Score 530.5; DB 2; Length 200;  
Best Local Similarity 55.8%; Pred. No. 3.8e-35;  
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDYDFKMKTIIEVDGKIVRI 59  
Db 1 MAKTYDHLFKLLIGDSGVGKTCVLFPSDAFNTTISTIGIDFKIKTVLHGKKIKL 60  
Qy 60 QIWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWVSDVDVEYAPGKILIG 119  
Db 61 QIWDTAGQERFHTITAYYRGAMGILVYDITNGKSFENISKWLNRNDEHANEDVERMLL 120  
Qy 120 GNKADEOKRVGREGQOOLAKYGMDFYETSACTNLNIKESFTRLTELVLQ 171  
Db 121 GNKCMDDKRVLPKGEQIAREHGIRFFETSAKVNINIEKAFTLTAEIDILR 172

## RESULT 8

B38625  
GTP-binding protein ora2 - electric ray (Discopyge ommata)  
C;Species: Discopyge ommata  
C;Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 02-Feb-2001  
C;Accession: B38625  
R;Ngsee, J.K.; Elferink, L.A.; Scheller, R.H.  
J. Biol. Chem. 266, 2675-2680, 1991  
A;Title: A family of ras-like GTP-binding proteins expressed in electromotor neurons.  
A;Reference number: A38625; MUID:91115900; PMID:1899244  
A;Accession: B38625  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-209 <NGS>  
A;Cross-references: GB:M38391  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;  
F;9-124/Domain: translation elongation factor Tu homology <ETU>  
F;15-22/Region: nucleotide-binding motif A (P-loop)  
F;121-124/Region: GTP-binding NKXD motif  
F;150-152/Region: GTP-binding SAK/L motif  
F;206/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 48.0%; Score 530; DB 2; Length 209;  
Best Local Similarity 47.4%; Pred. No. 4.4e-35;  
Matches 99; Conservative 56; Mismatches 44; Indels 10; Gaps 3;  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEPHSSHISTIGVDYDFKMKTIIEVDGKIVRIQ 60  
Db 1 MAKTYDHLFKLLIGDSGVGKTCVLFPSDAFNTTISTIGIDFKIKTVLHGKKIKLQ 60  
Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWVSDVDVEYAPGKILIG 120  
Db 61 IWDTAGQERFHTITAYYRGAMGILVYDITNEKSFENIKWIRNIEHASSDVERMLG 120  
Qy 121 NKADDEOKRVGREGQOOLAKYGMDFYETSACTNLNIKESFTRLTELVLQAHKKELEGL 180  
Db 121 NKCDMNEKRVSKERGEKLAIDYGIKF-ETSAKSSINVEEAFITLARDIMTKLNKQM--- 176



A;Residues: 1-201 <WAM>  
A;Cross-references: EMBL:U80033; PIDN:AA48200.1; GSPDB:GN00019; CESP:T23H2.5  
A;Experimental source: strain Bristol N2; clone T23H2  
C;Genetics:  
A;Gene: CESP:T23H2.5  
A;Map position: 1  
A;Introns: 43/1; 147/2; 173/3  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 46.5%; Score 514; DB 2; Length 201;  
Best Local Similarity 46.9%; Pred. No. 8e-34;  
Matches 100; Conservative 48; Mismatches 51; Indels 14; Gaps 3;

Qy 1 MAKQ-YDVLFRLLIGSGVGKTCILCRFTDNEFHSHSTIGVDFKMTIEVDGKVRIOIWD 59  
Db 1 MARRPYDMLFKLLIGSGVGKTCILYRSDDAFTTFTSTIGIDFKIRKILGKKIKL 60

Qy 60 QIWDTAGQERYQTITKQYRRAAGIFLVYDISSERSYQHIMKWVSDVDEYAPGQVQKILI 119  
Db 61 QIWDTAGQERFHTITTSYRGAMGIMLVYDITNAKSFDTNAKWLRTIDEHASEDVVKMIL 120

Qy 120 GNKADERQKRVGREGQOLAKYGMDFYETSACTNLNIKESFTRLTVELVLQAHRKLE 179  
Db 121 GNKCDMSDRVRSRGEKIAQDHGISHFETSAKLNVRVDTAFYDLAEAIL----- 171

Qy 180 LRMRASNELALAELEEBEGKEGPEGPANSSKTCWC 212  
Db 172 AKMPDSTD-----EQSRDVTNVPVQVQSSSGGC 200

RESULT 13  
S33900  
GTP-binding protein ypt2 - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 02-Feb-2001  
C;Accession: S33900; JQ2233  
R;Fleming, A.J.; Mandel, T.; Roth, I.; Kuhlmeier, C.  
Plant Cell 5, 297-309, 1993  
A;Title: The patterns of gene expression in the tomato shoot apical meristem.  
A;Reference number: S33899; MUID:93222691; PMID:8467223  
A;Accession: S33900  
A;Molecule type: mRNA  
A;Residues: 1-216 <FL2>  
A;Cross-references: GB:X69980; NID:g313028; PIDN:CAA49600.1; PID:g313029  
C;Genetics:  
A;Gene: ypt2  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
C;Keywords: GTP binding; membrane protein; nucleotide binding; P-loop  
F;16-131/Domain: translation elongation factor Tu homology <ETU>  
F;22-29/Region: nucleotide-binding motif A (P-loop)  
F;128-131/Region: GTP-binding NKXD motif  
F;159-161/Region: GTP-binding SAK/L motif

Query Match 46.3%; Score 512; DB 2; Length 216;  
Best Local Similarity 50.5%; Pred. No. 1.3e-33;  
Matches 99; Conservative 41; Mismatches 44; Indels 12; Gaps 3;

Qy 5 YDVLFRLLIGSGVGKTCILCRFTDNEFHSHSTIGVDFKMTIEVDGKVRIOIWD 64  
Db 12 YDYLKLLIGDGTGVGKSCLLRFSDFSTFTTIGIDFKIRTIYELDGKRIKLQIWD 71

Qy 65 AGQERYOTITKQYRRAAGIFLVYDISSERSYQHIMKWVSDVDEYAPGQVQKILGNKAD 124  
Db 72 AQGERFRITTYAYRGAMGILLVYDVTDESFNIRNWKNIHQHSDNVNKLIVGNKAD 131

Qy 125 -EEQKRVGREGQOLAKYGMDFYETSACTNLNIKESF-----TRLTVELVLQAHRK 175  
Db 132 MDESKRAVPTSKGQALADEYGIKFFETSAKTNLNVNPFVMSIAKDQIKRLTETDTKA 188

Qy 176 ELEGLRMRASNELALAELE 191  
Db 189 EPQGITRINQSDQAGTA 204

RESULT 14  
T48378  
GTP-binding protein-like - Arabidopsis thaliana  
N;Alternate names: protein F12E4.300  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 02-Sep-2000  
C;Accession: T48378  
R;Bevan, M.; Pohl, T.; Weizensegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
submitted to the Protein Sequence Database, March 2000  
A;Reference number: Z24492  
A;Accession: T48378  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-216 <BEV>  
A;Cross-references: EMBL:AL162751  
A;Experimental source: cultivar Columbia; BAC clone F12E4  
C;Genetics:  
A;Map position: 5  
A;Introns: 25/1; 49/2; 83/1; 105/1; 137/3; 158/3; 189/3  
A;Note: F12E4.300  
C;Superfamily: ras transforming protein; translation elongation factor Tu homology  
F;16-131/Domain: translation elongation factor Tu homology <ETU>

Query Match 46.3%; Score 512; DB 2; Length 216;  
Best Local Similarity 49.8%; Pred. No. 1.3e-33;  
Matches 100; Conservative 43; Mismatches 46; Indels 12; Gaps 3;

Qy 5 YDVLFRLLIGSGVGKTCILCRFTDNEFHSHSTIGVDFKMTIEVDGKVRIOIWD 64  
Db 12 YDYLKLLIGDGTGVGKSCLLRFSDFSTFTTIGIDFKIRTIYELDGKRIKLQIWD 71

Qy 65 AGQERYOTITKQYRRAAGIFLVYDISSERSYQHIMKWVSDVDEYAPGQVQKILGNKAD 124  
Db 72 AQGERFRITTYAYRGAMGILLVYDVTDESFNIRNWKNIHQHSDNVNKLIVGNKAD 131

Qy 125 -EEQKRVGREGQOLAKYGMDFYETSACTNLNIKESF-----TRLTVELVLQAHRK 175  
Db 132 MDESKRAVPTAKGQALADEYGIKFFETSAKTNLNVNPFVMSIAKDQIKRLTETDTKA 188

Qy 176 ELEGLRMRASNELALAELE 196  
Db 189 EPQGITKITQDTAASSSTAER 209

RESULT 15  
S12790  
GTP-binding protein ypt2 - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 02-Feb-2001  
C;Accession: S12790; S10493; T39214  
R;Haubruck, H.; Engelke, U.; Mertins, P.; Gallwitz, D.  
EMBO J. 9, 1957-1962, 1990  
A;Title: Structural and functional analysis of ypt2, an essential ras-related gene in the  
fission yeast Schizosaccharomyces pombe.  
A;Reference number: S12790; MUID:90269232; PMID:2112089  
A;Accession: S12790  
A;Molecule type: DNA  
A;Residues: 1-200 <HAU>  
A;Cross-references: EMBL:X52469; NID:g5143; PIDN:CAA36707.1; PID:g5144  
R;Fawell, E.; Hook, S.; Sweet, D.; Armstrong, J.  
Nucleic Acids Res. 18, 4264, 1990  
A;Title: Novel YPT1-related genes from Schizosaccharomyces pombe.  
A;Reference number: S10492; MUID:90332438; PMID:2115995  
A;Accession: S10493  
A;Molecule type: DNA  
A;Residues: 1-200 <FAW>  
A;Cross-references: EMBL:X52864; NID:g5149; PIDN:CAA37045.1; PID:g5150  
R;McDougall, R.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997  
A;Reference number: Z21836  
A;Accession: T39214  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

```

A;Residues: 1-200 <MCD>
A;Cross-references: EMBL:Z99262; PIDN:CA816405.1; GSPDB:GN00066; SPDB:SPAC39E9.07C
A;Experimental source: strain 972h-; cosmid c9E9
C;Genetics:
A;Gene: ypt2
A;Map position: 1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;
F;10-125/Domain: translation elongation factor Tu homology <TU>
F;16-23/Region: nucleotide-binding motif A (P-loop)
F;122-125/Region: GTP-binding NKXD motif
F;152-154/Region: GTP-binding SAK/L motif
F;199,200/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match      46.2%; Score 510; DB 2; Length 200;
Best Local Similarity 51.9%; Pred. No. 1.6e-33;
Matches 97; Conservative 39; Mismatches 43; Indels 8; Gaps 1;

QY 3 KQYDVLPRLLIGDSGVGKTCLLCRFTDNEPSSHISTIGVDYFKMKTIEVDGKIVRIQIW 62
   |||||
Db 4 KSYDYLKLLIGDSGVGKSCLLRFSEDSFTPSFITITIGIDFKIRTELGGKIKLQIW 63
   |||||

QY 63 DTAGQERYQITIKQYRRAGQIFVLVYDISSRSYQHIMKWVSDVDEYAPGQVKILGNK 122
   |||||
Db 64 DTAGQERFRITITAYYRGAMGILLYDVTDKKSPDNVTRTFPSNVEQHASENVYKILGNK 123
   |||||

QY 123 ADREKQKVGRECCQQLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHRKELEGLRM 182
   |||||
Db 124 CDCEQKQVSECCQQLADELGVKFLASAKTNVNDPEAFTLA-----REIKKQKI 175
   |||||

QY 183 RASNELA 189
   |||||
Db 176 DAENEPS 182

```

Search completed: March 15, 2004, 11:16:52  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:09:18 ; Search time 18 Seconds  
(without alignments)  
613.271 Million cell updates/sec

Title: US-09-817-198c-2

Perfect score: 1105

Sequence: 1 MARQYDVLFRLLIGDSGVG.....LEEEKGKPGPANSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1105	100.0	212	1 RB15_HUMAN	P59190 homo sapien
2	1090	98.6	212	1 RB15_MOUSE	Q9K386 mus musculus
3	1077	97.5	212	1 RB15_RAT	P35289 rattus norv
4	545	49.3	207	1 RB8B_HUMAN	P24407 homo sapien
5	539.5	48.8	210	1 RB8B_DISOM	P22128 discopyge o
6	532	48.1	203	1 RB13_HUMAN	P31153 homo sapien
7	531.5	48.1	200	1 RB10_DISOM	P22127 discopyge o
8	530.5	48.0	200	1 RB10_CANFA	P24409 canis famil
9	530.5	48.0	200	1 RB10_HUMAN	O88386 homo sapien
10	528.5	47.8	207	1 RB8B_HUMAN	Q92930 homo sapien
11	528.5	47.8	207	1 RB8B_RAT	P70550 rattus norv
12	526	47.6	206	1 RB8B_MOUSE	P55258 mus musculu
13	510	46.2	200	1 YPT2_SCHPO	P17609 schizosacch
14	509	46.1	215	1 RB11_BETVU	Q93433 beta vulgar
15	508	46.0	216	1 AR33_ARATH	P28186 arabidopsis
16	505.5	45.7	217	1 YPT2_VOLCA	P36861 volvox cart
17	502.5	45.5	200	1 RB10_RAT	P35281 rattus norv
18	502	45.4	203	1 RY11_YARLI	P41924 yarrowia li
19	493.5	43.8	208	1 SAS1_DICDI	P20790 dictyosteli
20	483	43.7	208	1 YPT1_MAYZE	P16976 zea mays (m
21	480	43.4	202	1 R1C1_ORVSA	P40392 oryza sativ
22	479	43.3	203	1 SAS2_DICDI	P20791 dictyosteli
23	478.5	43.3	201	1 RB1B_HUMAN	Q9H044 homo sapien
24	477.5	43.2	201	1 RB35_HUMAN	Q15286 homo sapien
25	477.5	43.2	203	1 YPT1_CHLRE	Q39571 chlamydomon
26	477	43.2	201	1 YPT1_PHYIN	Q01890 phytophor
27	476	43.1	258	1 AR45_ARATH	P28188 arabidopsis
28	475.5	43.0	203	1 YPT1_VOLCA	P31584 volvox cart
29	475	43.0	203	1 YPT1_NEUCR	P33723 neurospora
30	473	42.8	210	1 SEC4_CANAL	O4462 candida alb
31	470.5	42.6	203	1 YPT2_MAYZE	Q05737 zea mays (m
32	470.5	42.6	215	1 SEC4_YEAST	P07560 saccharomyc
33	469	42.4	201	1 RB1B_RAT	P10536 rattus norv

#### RESULT 1

ID	RB15_HUMAN	STANDARD;	PRT;	212 AA.
AC	P59190;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Ras-related protein Rab-15.			
GN	RAB15.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OK	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=12508121;			
RA	Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,			
RA	Cattolico L., Levy M., Barbe V., De Berardinis V., Ureta-Vidal A.,			
RA	Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S.,			
RA	Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,			
RA	Bruels T., Jaillon O., Friedlander L., Samson G., Brottier P.,			
RA	Cure S., Segurens B., Aniere F., Samain S., Crespeau H., Abbasi N.,			
RA	Aiach N., Boscut D., Dickhoff R., Dors M., Dubois I., Friedman C.,			
RA	Gouyvenoux M., James R., Madan A., Mairey-Estrada B., Mangelot S.,			
RA	Martins N., Menard M., Oztas S., Ratcliffe A., Shaffer T., Trask B.,			
RA	Vacherie B., Bellemere C., Belser C., Besnard-Gonnet M.,			
RA	Bartol-Mavel D., Boutard M., Briez-Silla S., Combette S.,			
RA	Dufosse-Laurent V., Ferron C., Lechaplais C., Louesse C., Muselet D.,			
RA	Magdelenat G., Pateau E., Petit E., Sirvain-Trukiewicz P., Trybou A.,			
RA	Vega-Czarny N., Bataille E., Bluet E., Bodelais I., Dubois M.,			
RA	Dumont C., Guerin T., Hafray S., Hammadi R., Muanga J., Pellouin V.,			
RA	Robert D., Wunderle E., Gauguet G., Roy A., Sainte-Marthe L.,			
RA	Verdier J., Verdier-Discala C., Hillier L.W., Fulton L., McPherson J.,			
RA	Matsuda F., Wilson R., Scarpelli C., Gvavay G., Wincker P., Saurin W.,			
RA	Quetier F., Waterston R., Hood L., Weissbach J.,			
RT	"The DNA sequence and analysis of human chromosome 14."			
RL	Nature 421:601-607(2003).			
CC	-!- FUNCTION: May act in concert with RAB3A in regulating aspects of synaptic vesicle membrane flow within the nerve terminal (By similarity).			
CC	-!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AL139022; -; NOT ANNOTATED\_CDS.  
Genew; HGNC:20150; RAB15.  
InterPro; IPR003579; GTPase\_Rab.  
InterPro; IPR003577; GTPase\_Ras.  
InterPro; IPR003578; GTPase\_Rho.  
InterPro; IPR002041; RAN.



```
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR SMART; SM00176; RAN; 1.
DR SMART; SM00173; RAS; 1.
DR SMART; SM00174; RHO; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT LIPID 210 210 S-geranylgeranyl cysteine
FT (By similarity).
FT LIPID 212 212 S-geranylgeranyl cysteine
FT (By similarity).
SQ SEQUENCE 212 AA; 24390 MW; D16A0C71797ED782 CRC64;

Query Match 100.0%; Score 1105; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 5e-81;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIKIEVDGKIRIQ 60
Db 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIKIEVDGKIRIQ 60

Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVDYDISSERSYOHIMKWSDVDYAPGQVKILIG 120
Db 61 IWDTAGQERYQTITKQYRRAQGIPLVDYDISSERSYOHIMKWSDVDYAPGQVKILIG 120

Qy 121 NKADEBKQKRVGREGQQLAKKEYGMDFYETSACTNLNLIKESFTRLTTLVLAQHRKELEGL 180
Db 121 NKADEBKQKRVGREGQQLAKKEYGMDFYETSACTNLNLIKESFTRLTTLVLAQHRKELEGL 180

Qy 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
Db 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212

RESULT 2
RB15 MOUSE
ID RB15 MOUSE STANDARD; PRT; 212 AA.
AC Q8K386;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras-related protein Rab-15.
GN RAB15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=23288257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buétow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerker A., Schein J.E., Jones S.J.M., Marra M.A.;
```

```
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: May act in concert with RAB1A in regulating aspects of
CC synaptic vesicle membrane flow within the nerve terminal (By
CC similarity).
CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; BC027769; AH27769.1; -.
CC MGD; MGI:1916865; 2310012G06Rik.
CC InterPro; IPR003579; GTPase_Rab.
CC InterPro; IPR003577; GTPase_Ras.
CC InterPro; IPR003578; GTPase_Rho.
CC InterPro; IPR002041; RAN.
CC InterPro; IPR001806; Ras trnsfrmg.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF00071; ras; 1.
CC PRINTS; PR00449; RASTRNSFRMG.
CC SMART; SM00175; RAB; 1.
CC SMART; SM00176; RAN; 1.
CC SMART; SM00173; RAS; 1.
CC SMART; SM00174; RHO; 1.
CC TIGRFAMs; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT LIPID 210 210 S-geranylgeranyl cysteine
FT (By similarity).
FT LIPID 212 212 S-geranylgeranyl cysteine
FT (By similarity).
SQ SEQUENCE 212 AA; 24318 MW; ESC492846DD47F12 CRC64;

Query Match 98.6%; Score 1090; DB 1; Length 212;
Best Local Similarity 98.1%; Pred. No. 7.9e-80;
Matches 208; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIKIEVDGKIRIQ 60
Db 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIKIEVDGKIRIQ 60

Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVDYDISSERSYOHIMKWSDVDYAPGQVKILIG 120
Db 61 IWDTAGQERYQTITKQYRRAQGIPLVDYDISSERSYOHIMKWSDVDYAPGQVKILIG 120

Qy 121 NKADEBKQKRVGREGQQLAKKEYGMDFYETSACTNLNLIKESFTRLTTLVLAQHRKELEGL 180
Db 121 NKADEBKQKRVGREGQQLAKKEYGMDFYETSACTNLNLIKESFTRLTTLVLAQHRKELEGL 180

Qy 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
Db 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212

RESULT 3
RB15 RAT
ID RB15 RAT STANDARD; PRT; 212 AA.
AC P35289;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras-related protein Rab-15.
GN RAB15.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=92210533; PubMed=1313420;  
 RA Eiferink L.A., Anzai K., Scheller R.H.;  
 RT "Rab15, a novel low molecular weight GTP-binding protein specifically  
 RT expressed in rat brain."  
 RL J. Biol. Chem. 267:5768-5775 (1992).  
 RN [2]  
 RP ERATUM.  
 RX MEDLINE=93054572; PubMed=1429617;  
 RA Eiferink L.A., Anzai K., Scheller R.H.;  
 RL J. Biol. Chem. 267:22693-22693 (1992).  
 CC -!- FUNCTION: May act in concert with RAB3A in regulating aspects of  
 CC synaptic vesicle membrane flow within the nerve terminal.  
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in neural tissues.  
 CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M83679; AAA41995.1; -.  
 DR PIR; F42148; F42148.  
 DR HSPSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRPFAMS; TIGR00231; small GTP; 1.  
 DR GTP-binding; Lipoprotein; Prenylation; Protein transport.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT LIPID 210 210 S-geranylgeranyl cysteine  
 FT (By similarity).  
 FT LIPID 212 212 S-geranylgeranyl cysteine  
 FT (By similarity).  
 SQ SEQUENCE 212 AA; 24283 MW; 04817DDA66CADE12 CRC64;  
 Query Match 97.5%; Score 1077; DB 1; Length 212;  
 Best Local Similarity 97.6%; Pred. No. 8.5e-79;  
 Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRPTDNEFHSSHSITIGVDFPKMTIEVDGKVRIG 60  
 DB 1 MAKQYDVLFRLLIGDSGVGKTCLLCRPTDNEFHSSHSITIGVDFPKMTIEVDGKVRIG 60  
 QY 61 IWDTAGQERYTITKQYRRAQGIPLVYDIDSSERSYHIMKWSVDVDEYAPGVQKILIG 120  
 DB 61 IWDTAGQERYTITKQYRRAQGIPLVYDIDSSERSYHIMKWSVDVDEYAPGVQKILIG 120  
 QY 121 NKADBEOKRVGREGOGQOLAKKEYGMDVETSACTNLNLIKESFTRLTLVLAQHRKELEGL 180  
 DB 121 NKADBEOKRVGREGOGQOLAKKEYGMDVETSACTNLNLIKESFTRLTLVLAQHRKELEGL 180  
 QY 181 RMRASNELALAELEBEKGKPEGPANSSKTCWC 212  
 DB 181 RTCASNELALAELEBEKGKPEGPANSSKTCWC 212  
 RESULT 4  
 RAB8 HUMAN  
 ID RAB8\_HUMAN  
 AC P24407;

DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ras-related protein Rab-8 (Rab-8A) (Oncogene c-mel).  
 GN MEL OR RAB8.  
 OS Homo sapiens (Human), and  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606, 9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RX MEDLINE=94124602; PubMed=8294494;  
 RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,  
 RA Tavittian A., Louvard D.;  
 RT "A small rab GTPase is distributed in cytoplasmic vesicles in non  
 RT polarized cells but colocalizes with the tight junction marker ZO-1  
 RT in polarized epithelial cells."  
 RL J. Cell Biol. 124:101-115 (1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RX MEDLINE=91360267; PubMed=1886711;  
 RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,  
 RA Johnson K.J.;  
 RT "The MEL gene: a new member of the RAB/YPT class of RAS-related  
 RT genes."  
 RL Oncogene 6:1347-1351 (1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Brain;  
 RX Puhl H.L. III, Ikeda S.R., Aronstam R.S.;  
 RT "cDNA clones of human proteins involved in signal transduction  
 RT sequenced by the Guthrie cDNA resource center ([www.cdna.org](http://www.cdna.org))."  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Skin;  
 RX MEDLINE=22389257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=C. familiaris; STRAIN=Cocker spaniel;  
 RX MEDLINE=91061765; PubMed=2123294;  
 RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;  
 RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell  
 RT line."  
 RL Mol. Cell Biol. 10:6578-6585 (1990).  
 CC -!- FUNCTION: May be involved in vesicular trafficking and  
 CC neurotransmitter release.  
 CC -!- SUBUNIT: Interacts with MAP4K2 and SYTL4 (By similarity).  
 CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL; X56741; CAA0065.1; --  
 EMBL; S53268; AAB19881.1; --  
 EMBL; AF498943; AAM21091.1; --  
 EMBL; BC002977; AAH02977.1; --  
 EMBL; X56385; CAB56776.1; --  
 PIR; B36364; B36364.  
 PIR; B49647; B49647.  
 HSP; P05713; 3RAB.  
 Genew; HGNC:7007; MEL.  
 MIN; 165040; --  
 InterPro; IPR003579; GTPase Rab.  
 InterPro; IPR001806; Ras trnsfrmg.  
 InterPro; IPR005225; Small\_GTP.  
 Pfam; PF00071; ras; 1.  
 PRINTS; PR00449; RASTRNSFRMG.  
 SMART; SM00175; RAB; 1.  
 TIGRFAMs; TIGR00231; small GTP; 1.  
 GTP-binding; Prenylation; Lipoprotein; Protein transport;  
 KW Proto-oncogene.  
 FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
 FT LIPID 204 204 S-geranylgeranyl cysteine  
 (By similarity).  
 FT CONFLICT 177 183 LEGNSPO -> WKATAP (IN REF. 2).  
 SQ SEQUENCE 207 AA; 23668 MW; AA52DBF54A2CD056 CRC64;

Query Match 49.3%; Score 545; DB 1; Length 207;  
 Best Local Similarity 52.9%; Pred. No. 1.6e-36;  
 Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;  
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKIVRIQ 60  
 DB 1 MAKTYDYLKLLIGDSGVGKTCCLFRFSEDAFNSTFISTIGIDFKIRTVELDGKRIKLQ 60  
 QY 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGVQKILIG 120  
 DB 61 IWDTAGQERFTITTYIRGAMGIMLVYDITNEKSFNIRNWIKNIEHSAADVVERMILG 120  
 QY 121 NKADQKQVGRGQOQLAKEYGMDFYETSACTNLNIKESFTLRLTLVLQAHKKELEGL 180  
 DB 121 NKCDVNDKQVSKERGEKLAIDYGIKFWETSAKANINVENAFPTLARDIKAKMDKKLEGN 180  
 QY 181 RMRASNE 187  
 DB 181 SPOGSNQ 187

RESULT 5  
 RAB8 DISOM STANDARD; PRT; 210 AA.  
 ID -RAB8 DISOM  
 AC P22128;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ras-related protein Rab-8 (ORA2).  
 OS Discoprye omata (Electric ray).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Squala; Hyposquala; Pristiogaster; Batoidae;  
 OC Torpediniformes; Narcinoidei; Narcinidae; Discopyge.  
 NCBI\_TaxID=7785;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Electric lobe;

MEDLINE=91115900; PubMed=1899244;  
 RA Ngeese J.K., Elferink L.A., Scheller R.H.;  
 RT "A family of ras-like GTP-binding proteins expressed in electromotor  
 RT neurons.";  
 RL J. Biol. Chem. 266:2675-2680 (1991).  
 CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL; M38391; AAA9232.1; --  
 DR PIR; B38625; B38625.  
 DR HSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 GTP-binding; Prenylation; Lipoprotein.  
 KW NP\_BIND 15 22 GTP (BY SIMILARITY).  
 FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
 FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
 FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
 FT LIPID 207 207 S-geranylgeranyl cysteine  
 (By similarity).  
 SQ SEQUENCE 210 AA; 24164 MW; 0F73EDB0DB8B9E8A CRC64;

Query Match 48.8%; Score 539.5; DB 1; Length 210;  
 Best Local Similarity 47.4%; Pred. No. 4.5e-36;  
 Matches 99; Conservative 56; Mismatches 45; Indels 9; Gaps 2;  
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKIVRIQ 60  
 DB 1 MAKTYDYLKLLIGDSGVGKTCCLFRFSEDAFNSTFISTIGIDFKIRTVELDGKRIKLQ 60  
 QY 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGVQKILIG 120  
 DB 61 IWDTAGQERFTITTYIRGAMGIMLVYDITNEKSFNIRNWIKNIEHSAADVVERMILG 120  
 QY 121 NKADQKQVGRGQOQLAKEYGMDFYETSACTNLNIKESFTLRLTLVLQAHKKELEGL 180  
 DB 121 NKCDVNDKQVSKERGEKLAIDYGIKFWETSAKANINVENAFPTLARDIKAKMDKKLEGN 180  
 QY 181 RMRASNE 208  
 DB 178 -----NENSLQEAVDKLPKPKPKKK 201

RESULT 6  
 RB13 HUMAN STANDARD; PRT; 203 AA.  
 ID -RB13 HUMAN  
 AC P51153;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ras-related protein Rab-13.  
 GN RAB13.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94124602; PubMed=8294494;  
 RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,  
 RA Tavittian A., Louvard D.;

RT "A small rab GTPase is distributed in cytoplasmic vesicles in non  
RT polarized cells but colocalizes with the tight junction marker ZO-1  
RT in polarized epithelial cells.";  
RL J. Cell Biol. 124:101-115(1994).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Brain;  
RC Puhl H.L. III,  
RA "cDNA clones of human proteins involved in signal transduction  
RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Placenta;  
RC MEDLINE=2388252; PubMed=12477932;  
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Could participate in polarized transport, in the  
CC assembly and/or the activity of tight junctions.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC TIGHT JUNCTIONS OR ASSOCIATED  
CC WITH VESICLES SCATTERED THROUGHOUT THE CYTOPLASM IN CELLS LACKING  
CC TIGHT JUNCTIONS.  
CC -!- TISSUE SPECIFICITY: Detected in several types of epithelia,  
CC including intestine, kidney, liver, and in endothelial cells.  
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X75593; CAA53266.1; -.  
DR EMBL; AF498948; AAM21096.1; -.  
DR EMBL; BC000799; AAH00799.1; -.  
DR PIR; A49647; A49647.  
DR HSSP; P01112; 1PLI.  
DR GENE; HGNC:9762; RAB13.  
DR MIM; 602672; -.  
DR GO; GO:0005923; C:tight junction; TAS.  
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; TAS.  
DR GO; GO:0007155; P:cell adhesion; TAS.  
DR GO; GO:0016192; P:vesicle-mediated transport; TAS.  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001806; Ras trnsfrmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMs; TIGR00231; small GTP; 1.  
DR GTP-binding; Lipoprotein; Prenylation.  
KW NP\_BIND 15 22 GTP (BY SIMILARITY).  
FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
FT

FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
FT DOMAIN 37 45 S-geranylgeranyl cysteine  
FT LIPID 200 200 (By similarity).  
SQ SEQUENCE 203 AA; 22774 MW; 141621CB998178DA CRC64;  
Query Match 48.1%; Score 532; DB 1; Length 203;  
Best local similarity 47.6%; Pred. No. 1.7e-35;  
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;  
QY 1 MAKQYDVLFRLLILGDSGVKTCCLLCPTDNEHSSHSITGVDFPKMTIEVDGKVRIO 60  
DB 1 MAKAYDHLFLKLLIGDSGVKTCCLIRPAEDNFNTTIGIDFKIRTVDIEGKKIKQ 60  
QY 61 IWTAGQRYQTITKQYRRAQGIIFLYDYSERSYQHIMKWSDVDVEYAPGVQKILIG 120  
DB 61 VWDTAGQERKTIITAYYRGAMGILLVYDITDEKSFENIQNWMSIKENASAGVERLLLG 120  
QY 121 NKADDEQKQVGRGQGOOLAKYGMDFYTSACTNLNKSFTLTLVLQAHKKEGL 180  
DB 121 NKCDMEAKRKVQKEQADKLAREHGIRFETSAKSSMMVDEAFSSLDAIL-----LKSG 174  
QY 181 RMRASNELALAELEEEKPEGPANSSKTC 210  
DB 175 GRRSGN-----GNKP--PSTDUKTC 192  
RESULT 7  
RB10\_DISOM STANDARD; PRT; 200 AA.  
AC P22127;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ras-related protein Rab-10 (ORA1).  
OS Discopoge ommata (Electric ray).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Squalia; Hypnosqualia; Pristioraja; Batoidae;  
OC Torpediniformes; Narcinoidei; Narcinidae; Discopoge.  
OX NCBI\_TaxID=7785;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Electric lobe;  
RX MEDLINE=91115900; PubMed=1899244;  
RA Ngeese J.K., Elferink L.A., Scheller R.H.;  
RT "A family of ras-like GTP-binding proteins expressed in electromotor  
RT neurons.";  
RL J. Biol. Chem. 266:2675-2680(1991).  
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M38390; AAA49230.1; -.  
DR PIR; A38625; A38625.  
DR HSSP; P05713; 3RAB.  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001806; Ras trnsfrmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMs; TIGR00231; small GTP; 1.  
DR GTP-binding; Lipoprotein; Prenylation.  
KW NP\_BIND 16 23 GTP (BY SIMILARITY).  
FT NP\_BIND 64 68 GTP (BY SIMILARITY).  
FT NP\_BIND 122 125 GTP (BY SIMILARITY).  
FT LIPID 198 198 S-geranylgeranyl cysteine

```
FT LIPID 199 199 (By similarity).
FT S-geranylgeranyl cysteine
FT (By similarity).
SQ SEQUENCE 200 AA; 22623 MW; 41D38E3D760519C5 CRC64;
Query Match 48.1%; Score 531.5; DB 1; Length 200;
Best Local Similarity 56.4%; Pred. No. 1.8e-35;
Matches 97; Conservative 44; Mismatches 30; Indels 1; Gaps 1;
QY 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKIVRI 59
Db 1 MAKTYDLLFKLLIGDSGVGKTCVLFPSDDAFNTFTSTIGIDFKIKTVLHGGKIKL 60
QY 60 QIWDTAGOERYQITTKOYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVQKILI 119
Db 61 QIWDTAGOERFHTTTSYYRGAMGIMLVYDITNAKSPENISKLRNIDEHANEDEVRL 120
QY 120 GNKADBEQKRVQREOQQLAKYGMDFYETSACTNLNLIKESFTRLTFLVLQ 171
Db 121 GNKCDMEDKRVVLKSKGEQIAREHAIRFFETSANINIEKAFITLAEIDILQ 172
RESULT 8
RB10 CANFA STANDARD; PRT; 200 AA.
AC P24409;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Cocker spaniel;
RX MEDLINE=91061765; PubMed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
line.";
RL Mol. Cell. Biol. 10:6578-6585(1990).
CC -1- FUNCTION: May be involved in vesicular trafficking and
CC neurotransmitter release.
CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X56387; CAA379798.1; -.
DR PIR; D36364; D36364.
DR HSRF; P05713; 3RAB.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PRQ0449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 199 199 S-geranylgeranyl cysteine
FT (By similarity).
FT LIPID 200 200 S-geranylgeranyl cysteine
```

```
FT SEQUENCE 200 AA; 22569 MW; 5D52B8E8E47D4362 CRC64;
Query Match 48.0%; Score 530.5; DB 1; Length 200;
Best Local Similarity 55.8%; Pred. No. 2.2e-35;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFPKMTIEVDGKIVRI 59
Db 1 MAKTYDLLFKLLIGDSGVGKTCVLFPSDDAFNTFTSTIGIDFKIKTVLHGGKIKL 60
QY 60 QIWDTAGOERYQITTKOYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVQKILI 119
Db 61 QIWDTAGOERFHTTTSYYRGAMGIMLVYDITNAKSPENISKLRNIDEHANEDEVRL 120
QY 120 GNKADBEQKRVQREOQQLAKYGMDFYETSACTNLNLIKESFTRLTFLVLQ 171
Db 121 GNKCDMEDKRVVLKSKGEQIAREHAIRFFETSANINIEKAFITLAEIDILR 172
RESULT 9
RB10 HUMAN STANDARD; PRT; 200 AA.
AC O88386; Q9D7X6;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Homo sapiens (Human), and
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 10090;
RN [1]
SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Pituitary;
RX MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
axis and full-length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
RN [2]
SEQUENCE FROM N.A.
RC SPECIES=Human;
RX Wong K., Hong W., Tang B.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RC SPECIES=Human;
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Brain;
RA Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT "cDNA clones of human proteins involved in signal transduction
sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Cervix;
```

RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Heieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Mouse;  
RA zeng Q., Tan Y.H., Hong W.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Mouse; STRAIN=C57BL/6J; TISSUE=Stomach;  
RX MEDLINE=2108560; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata Y., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaado I., Pesole G., Quackenbush J.,  
RA Schirni L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyoko-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
CC -!- FUNCTION: May be involved in vesicular trafficking and  
CC neurotransmitter release.  
CC  
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF106681; AAD43034.1; -  
DR EMBL; AF297660; AAG13413.1; -  
DR EMBL; AF423223; BAB14474.1; -  
DR EMBL; AF498945; AAM21093.1; -  
DR EMBL; BC000896; AAH00896.1; -  
DR EMBL; AF035646; AAC29313.1; -  
DR EMBL; AK008725; BAB25858.1; -  
DR HSP; P05713; 3RAB;  
DR Genew; HGNC:9759; RAB10.  
DR MGD; MGI:105066; RAB10.  
DR InterPro; IPR003579; GTPase\_Rab.

DR InterPro; IPR001806; Ras\_trnsmfmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; Ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMNG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMs; TIGR00231; small GTP; 1.  
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.  
FT NP\_BIND 16 23 GTP (BY SIMILARITY).  
FT NP\_BIND 64 68 GTP (BY SIMILARITY).  
FT NP\_BIND 122 125 GTP (BY SIMILARITY).  
FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 199 199 S-geranylgeranyl cysteine  
FT (By similarity).  
FT LIPID 200 200 S-geranylgeranyl cysteine  
FT (By similarity).  
FT CONFLICT 106 106 N -> H (IN REF. 7).  
FT SEQUENCE 200 AA; 22541 MW; 7F02B8E8E46E1E8 CRC64;  
Query Match 48.0%; Score 530.5; DB 1; Length 200;  
Best Local Similarity 55.8%; Pred. No. 2.2e-35;  
Matches 96; Conservative 29; Mismatches 29; Indels 1; Gaps 1;  
Qy 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDPMKTTIEVDGIKVI 59  
Db 1 MAKTYDLLFKLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDPMKTTIEVDGIKVI 60  
Qy 60 QIWDTAGQERYQITTKOYVRAQGIFLVYDISSRSYOHIMKWYSDVDVYAPGQKILI 119  
Db 61 QIWDTAGQERPHITTSYRGMGIMLVYDITNGKSFENISKWRNIDEHANEDVERMLL 120  
Qy 120 GNKADEBQKRGVQEQQLAKYGMDFYTSACTNLNLIKESFRLTFLVLIQ 171  
Db 121 GNKCDMDKRVVPKKGQIAREHGIRFFETSANINIEKAFITLAEDILR 172  
RESULT 10  
R88B HUMAN  
ID R88B HUMAN STANDARD; PRT; 207 AA.  
AC Q92930; Q92933;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ras-related protein Rab-8B.  
GN RAB8B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RA Seki N., Saito T.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Heieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,

```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP [3]
RQ SEQUENCE OF 67-119 FROM N.A.
RX MEDLINE=96393028; PubMed=9799816;
RZ TISSUE=Melanoma;
RA Chen D., Guo J., Gahl W.A.;
RT "RAB GTPases expressed in human melanoma cells.";
RL Biochim. Biophys. Acta 1355:1-6 (1997).
CC -!- FUNCTION: May be involved in vesicular trafficking and
CC neurotransmitter release (By similarity).
CC -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB038995; BAA92249.1; -.
DR EMBL; BC020654; AAH20654.1; -.
DR EMBL; U66624; AAC51199.1; -.
DR HSP; P05713; 3RAB.
DR GO; GO:0003928; P:RAB small monomeric GTPase activity; NAS.
DR GO; GO:0015031; P:protein transport; NAS.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
DR GTP-binding; Prenylation; Lipoprotein; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 204 204 S-geranylgeranyl cysteine
FT (By similarity).
SQ SEQUENCE 207 AA; 23584 MW; 5960993C0F87F944 CRC64;
Query Match 47.8%; Score 528.5; DB 1; Length 207;
Best Local Similarity 54.0%; Pred. No. 3.3e-35;
Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;
Qy 1 MAKQYDLVFLRLLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFKMKTIEDVGKVIQ 60
Db 1 MAKTYDVLKLLIGDSGVGKTCLLFRSEDAFNTFTISTIGIDFKRTIELDGKKIKLQ 60
Qy 61 IWDTAGERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVDEYAPGQVKILIG 120
Db 61 IWDTAGOERFTITTYRGAAGMLVYDITNEKSPDNKNWIRNIEHASSDVERMILG 120
Qy 121 NKADEOKRQVREGQOOLAKYGMDFVETSACTNLNLIKES-FTRLTELVLQAHK 175
Db 121 NKCDMDNRQVSKERGEKLAIDYGIKLETSKSNVVEAFFTLARDIMTKLNK 176
Qy 121 NKDEEOKRQVREGQOOLAKYGMDFVETSACTNLNLIKES-FTRLTELVLQAHK 175
Db 121 NKCDMDNRQVSKERGEKLAIDYGIKLETSKSNVVEAFFTLARDIMTKLNK 176
RESULT 11
RB8B RAT STANDARD; PRT; 207 AA.
AC P70550;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras-related protein Rab-8B.
GN RAB8B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Muridae; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.

```



XX MEDLINE=91360267; PubMed=1886711;  
RA Nimm E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,  
RA Johnson K.J.;  
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related  
RT genes.";  
RL Oncogene 6:1347-1351(1991).  
RN [2]  
RP INTERACTION WITH MAP4K2.  
RC STRAIN=BALB/C; Tissue=Melanoma;  
RC MEDLINE=96209873; PubMed=8643544;  
RA Ren M., Zeng J., De Lemos-Charandini C., Rosenfeld M., Adenik M.,  
RA Sabatini D.D.;  
RT "In its active form, the GTP-binding protein rab8 interacts with a  
RT stress-activated protein kinase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:5151-5155(1996).  
RN [3]  
RP INTERACTION WITH SYTL4.  
RC MEDLINE=22590467; PubMed=12590134;  
RA Fukuda M.;  
RT "Glp4-a/granophilin-a inhibits dense-core vesicle exocytosis through  
RT interaction with the GTP-bound form of Rab27A in PC12 cells.";  
RL J. Biol. Chem. 278:15390-15396(2003).  
CC -1- FUNCTION: May be involved in vesicular trafficking and  
CC neurotransmitter release.  
CC -1- SUBUNIT: Interacts with MAP4K2 and SYTL4.  
CC -1- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; S53270; AB19682.1; -  
DR HSPSP; P05713; 3RAB.  
DR MGD; MG1:96960; Mel.  
DR GO; GO:0005515; F:protein binding; IPI.  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR001806; Ras\_trnsfrmg.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMs; TIGR00231; small GTP; 1.  
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;  
KW Proto-oncogene.  
FT NP\_BIND 15 22 GTP (BY SIMILARITY).  
FT NP\_BIND 63 67 GTP (BY SIMILARITY).  
FT NP\_BIND 121 124 GTP (BY SIMILARITY).  
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).  
FT LIPID 203 203 S-geranylgeranyl cysteine  
FT (By similarity).  
SQ SEQUENCE 206 AA; 23557 MW; 49D832725D62942 CRC64;  
Query Match 47.6%; Score 526; DB 1; Length 206;  
Best Local Similarity 56.4%; Pred. No. 5.2e-35;  
Matches 93; Conservative 41; Mismatches 31; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLLDGSGVGTCLLCRFTDNEPHSSHSITGIDVDFPKMTIEVDGKIRIQ 60  
Db 1 MAKYDYVLFKLLLDGSGVGTCLVFRSEDAFNSTFTSTIGIDFKRTIELDGKRIKIQ 60  
Qy 61 IWDTAGERYTITKQYRRAGQIFLVYDISSERSYQIMKWKVSDVEYAPGVQKILIG 120  
Db 61 IWDTAGQERFTITATYRGAMGLVYDITNEKSFQIRNWRINRIBEHASADVVERKILG 120  
Qy 121 NKABEQRQVGRGQQLAKYEGMDVYETSACTNLNKKESFTPL 165  
Db 121 NKDVDNDKQVSRGEGKALDYGIKFMETSAKANINVENAFPTL 165

RESULT 13  
YPT2 SCHPO  
ID YPT2 SCHPO STANDARD; PRT; 200 AA.  
AC P17609;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ras-related protein ypt2.  
GN YPT2 OR SPAC9E9.07C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RC MEDLINE=90269232; PubMed=2112089;  
RA Hengst L., Lehmeier T., Gallwitz D.;  
RT "Structural and functional analysis of ypt2, an essential ras-related  
RT gene in the fission yeast Schizosaccharomyces pombe encoding a Sec4  
RT protein homologue.";  
RL EMBL J. 9:1957-1962(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RC MEDLINE=90332438; PubMed=2115995;  
RA Fawell E., Hook S., Sweet D., Armstrong J.;  
RT "Novel YPT1-related genes from Schizosaccharomyces pombe.";  
RL Nucleic Acids Res. 18:4264-4264(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RC MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Skelton J., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Art R., Robben J., Grynoprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC -1- FUNCTION: Protein transport. Probably involved in vesicular  
CC traffic (By similarity).  
CC -1- MISCELLANEOUS: This protein is essential for cell viability.  
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.  
CC PROBABLE YEAST SEC4 HOMOLOG.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

EMBL; X52469; CAA36707.1; --  
 DR EMBL; X52864; CAA37045.1; --  
 DR EMBL; X59262; CAB16405.1; --  
 DR PIR; S12790; S12790.  
 DR HSSP; P05713; 3RAB.  
 DR GeneDB SPombe; SPAC9E9.07c; --  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport.  
 FT NP\_BIND 16 23 GTP (BY SIMILARITY).  
 FT NP\_BIND 64 68 GTP (BY SIMILARITY).  
 FT NP\_BIND 122 125 GTP (BY SIMILARITY).  
 FT DOMAIN 38 46 EFFECTOR REGION (PROBABLE).  
 FT LIPID 199 199 S-geranylgeranyl cysteine  
 FT (By similarity).  
 FT LIPID 200 200 S-geranylgeranyl cysteine  
 FT (By similarity).  
 SQ SEQUENCE 200 AA; 22755 MW; 2C658D153A290C30 CRC64;  
 Query Match 46.2%; Score 510; DB 1; Length 200;  
 Best Local Similarity 51.9%; Pred. No. 9.5e-34;  
 Matches 97; Conservative 39; Mismatches 43; Indels 8; Gaps 1;  
 QY 3 KOYDVLRLLLIGDSGVGKTCLLCRFTDNEPHSHSHSTIGVDPRKMTIEVDGKVRQIW 62  
 DB 4 KSYDYLKLLIGDSGVGKSCLLRFSDFSTPTSFITIGIDFKIRTIELDKRIKLQIW 63  
 QY 63 DTAGQRYVOTTKQYVRAQGIPLVYDISSERSYOHIMKWSVDVEYAPGVQKILGNK 122  
 DB 64 DTAGQRFRTITAYRGAMGILLVDYDKSPDNVRTWFSNVEQHASENVYKILIGNK 123  
 QY 123 ADEKQKQVGEQOQOLAKGYMDFYTSACTNLNLIKESFRLTELVLQAHKKELEGLRM 182  
 DB 124 CDCEQKQVFEQOQALDELGVKFLASAKTNVNDVEAFFLA-----REIKKQKI 175  
 QY 183 RASNELA 189  
 DB 176 DAENEFS 182  
 RESULT 14  
 RAB1\_BETVU  
 ID RAB1\_BETVU STANDARD; PRT; 215 AA.  
 AC Q39433;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ras-related protein RAB1B.  
 GN RAB1B.  
 OS Beta vulgaris (Sugar beet).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Amaranthaceae; Beta.  
 OX NCBI\_TaxID=161934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. D100 KS 38080;  
 RX MEDLINE=96307523; PubMed=8680960;  
 RA Dallery E., Quief S., Ben Jilany K.E., Kerckaert J.-P., Hagege D.;  
 RT "Molecular cloning and structural analysis of cDNAs that encode 3  
 small GTP-binding proteins from sugar beet."  
 CC C. R. Acad. Sci., III, Sci. Vie 319:91-97(1996).  
 RL -!- SIMILARITY: Belongs to the small GTPase superfamily. Rab family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 -----  
 DR EMBL; Z49152; CAA89021.1; --  
 DR PIR; T14565; T14565.  
 DR HSSP; P05713; 3RAB.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 KW GTP-binding; Lipoprotein; Prenylation.  
 FT NP\_BIND 22 29 GTP (BY SIMILARITY).  
 FT NP\_BIND 70 74 GTP (BY SIMILARITY).  
 FT NP\_BIND 128 131 GTP (BY SIMILARITY).  
 FT LIPID 212 212 S-geranylgeranyl cysteine  
 FT (By similarity).  
 FT LIPID 213 213 S-geranylgeranyl cysteine  
 FT (By similarity).  
 SQ SEQUENCE 215 AA; 23787 MW; 781CC5A17F29B8AA CRC64;  
 Query Match 46.1%; Score 509; DB 1; Length 215;  
 Best Local Similarity 47.8%; Pred. No. 1.2e-33;  
 Matches 100; Conservative 42; Mismatches 59; Indels 8; Gaps 3;  
 QY 5 YDVLRLLLIGDSGVGKTCLLCRFTDNEPHSHSHSTIGVDPRKMTIEVDGKVRQIW 64  
 DB 12 YDYLKLLIGDSGVGKSCLLRFSDFSTPTSFITIGIDFKIRTIELDKRIKLQIW 71  
 QY 65 AGQERYVOTTKQYVRAQGIPLVYDISSERSYOHIMKWSVDVEYAPGVQKILGNK 124  
 DB 72 AGQERFRTITAYRGAMGILLVDYDSEFNIRNIRNIHQHASDNDVKNKILVGNK 131  
 QY 125 -EOKRGVGEQOQOLAKGYMDFYTSACTNLNLIKESFRLTELVLQAHKKELEGLRM 183  
 DB 132 MDESKRAVPTAKQALADEYGIKFFTSAKTNLNVVEVFFSIARDIKQ--RLADSDTRQE 189  
 QY 184 ASNELALAELEEEEGKPEGPANSSKTCWC 212  
 DB 190 AQPSTITIKPADQ-----SGNQAAAKSACC 213  
 RESULT 15  
 ARAB3\_ARATH  
 ID ARAB3\_ARATH STANDARD; PRT; 216 AA.  
 AC P28186;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ras-related protein ARAB3.  
 GN ARAB3 OR AT3G46060 OR F12M12 30.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Landsberg erecta, cv. Columbia, cv. En-1, cv. Lapalmar,  
 RC and cv. Estland;  
 RC TISSUE=Leaf;  
 RX MEDLINE=92084144; PubMed=1748311;  
 RA Anai T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,  
 RA Matsui M.;  
 RT "Isolation and analysis of cDNAs encoding small GTP-binding proteins  
 of Arabidopsis thaliana."  
 RL Gene 108:259-264(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 11:12:33 ; Search time 46 Seconds  
(without alignments)  
1454.127 Million cell updates/sec

Title: US-09-817-198C-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798	72.2	168	11 Q91YW0	Q91yw0 mus musculus
2	573	51.9	211	4 Q8IW89	Q8iw89 homo sapien
3	563.5	51.0	200	13 Q7T3A4	Q7t3a4 brachydanio
4	557	50.4	211	5 Q9TYS2	Q9tys2 caenorhabdi
5	546.5	49.5	204	5 O15971	O15971 drosophila
6	541	49.0	207	11 Q8VCF6	Q8vcf6 mus musculus
7	528.5	47.8	203	11 Q8K3X5	Q8k3x5 rattus norv
8	528.5	47.8	207	5 O18338	O18338 drosophila
9	526.5	47.6	200	4 Q9H0T3	Q9h0t3 homo sapien
10	526	47.6	202	11 Q9DD03	Q9dd03 mus musculus
11	521	47.1	206	3 Q9HET4	Q9het4 aspergillus
12	519	47.0	216	10 Q24466	Q24466 arabidopsis
13	518	46.9	214	10 Q40218	Q40218 lotus japon
14	518	46.9	215	10 Q7XHP7	Q7xhp7 oryza sativ
15	515.5	46.7	216	10 Q40215	Q40215 lotus japon
16	515	46.6	216	10 Q9FUF1	Q9fuf1 arabidopsis

17	514	46.5	201	5 Q94148	Q94148 caenorhabdi
18	513.5	46.5	215	10 Q40219	Q40219 lotus japon
19	513	46.4	216	10 Q8VWF9	Q8vwf9 nicotiana t
20	512.5	46.4	214	10 Q84TA8	Q84ta8 oryza sativ
21	512	46.3	216	10 Q40177	Q40177 lycopersico
22	512	46.3	216	10 Q9LZD4	Q9lzd4 arabidopsis
23	511.5	46.3	212	10 Q40217	Q40217 lotus japon
24	510.5	46.2	216	10 Q8W3J4	Q8w3j4 nicotiana t
25	510	46.2	218	10 Q9SF91	Q9sf91 arabidopsis
26	508.5	46.0	215	10 Q41023	Q41023 pisum sativ
27	508	46.0	216	10 Q49844	Q49844 daucus caro
28	508	46.0	216	10 Q8W3J3	Q8w3j3 nicotiana t
29	507.5	45.9	222	10 Q96362	Q96362 brassica ca
30	507	45.9	202	3 Q96VL3	Q96vl3 collettotric
31	507	45.9	216	10 Q8LBO4	Q8lbo4 arabidopsis
32	506.5	45.8	216	10 Q9SWV8	Q9swv8 lycopersico
33	506	45.8	216	10 Q41024	Q41024 pisum sativ
34	505	45.7	198	4 Q96GU4	Q96gu4 homo sapien
35	505	45.7	216	10 Q8W3J2	Q8w3j2 nicotiana t
36	502.5	45.5	215	10 Q41022	Q41022 pisum sativ
37	495.5	44.8	204	3 Q9C1Z5	Q9c1z5 pichia past
38	492.5	44.6	215	10 Q41061	Q41061 pisum sativ
39	484	43.8	205	5 Q9BLF3	Q9blf3 entamoeba h
40	482	43.6	203	10 Q40569	Q40569 nicotiana t
41	482	43.6	205	5 Q26554	Q26554 schistosoma
42	481	43.5	203	10 Q8RU63	Q8ru63 oryza sativ
43	478	43.3	202	10 Q9ZRH6	Q9zrh6 petunia hyb
44	477	43.2	202	3 Q9HDT5	Q9hdt5 trichoderma
45	477	43.2	202	10 Q39845	Q39845 glycine max

#### ALIGNMENTS

#### RESULT 1

Q91YW0 ID Q91YW0 PRELIMINARY; PRT; 168 AA.  
AC Q91YW0;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN 2310012G06RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strauberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
DR EMBL; BC013790; A013790.1; -.  
DR HSSP; P05713; 3RAB.  
DR MGD; MGI:1916865; 2310012G06RIK.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
DR GO; GO:0015031; F:protein transport; IEA.  
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.  
DR InterPro; IPR003579; GTPase Rab.  
DR InterPro; IPR001806; Ras trnsfrmg.  
DR InterPro; IPR002078; Sig54\_interact.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRfams; TIGR00231; small GTP; 1.  
DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; 1.  
KW Hypothetical protein; GTP-binding; Lipoprotein.  
SQ SEQUENCE 168 AA; 19399 MW; 4E2C2FB1C56BCDF8 CRC64;

Query Match 72.2%; Score 798; DB 11; Length 168;  
Best Local Similarity 99.3%; Pred. No. 5.2e-60;  
Matches 152; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKIRIQ 60  
Db |||||  
Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIDVDGKIRIQ 60  
Db |||||  
Qy 61 IWDTAGERYQTITKQYRRAGGIFLVYDIDISSERSYQIMKWSVDVDEYAPGQKILIG 120  
Db |||||  
Qy 61 IWDTAGERYQTITKQYRRAGGIFLVYDIDISSERSYQIMKWSVDVDEYAPGQKILIG 120  
Db |||||  
Qy 121 NKADEEQKQVREGGQOLAKKEYGMDFYETSAC 153  
Db |||||  
Qy 121 NKADEEQKQVREGGQOLAKKEYGMDFYETSAC 153  
Db |||||

## RESULT 2

Q81W89 PRELIMINARY; PRT; 211 AA.  
AC Q81W89;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein (fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN SEQUENCE FROM N.A.  
RP TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC040679; AAH40679.1; -.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
DR GO; GO:0003930; F:RAS small monomeric GTPase activity; IEA.  
DR GO; GO:0003931; F:Rho small monomeric GTPase activity; IEA.  
DR GO; GO:0006886; P:intracellular protein transport; IEA.  
DR GO; GO:0007284; P:small GTPase mediated signal transduction; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. .; IEA.  
DR InterPro; IPR003579; GTPase\_Rab.  
DR InterPro; IPR003577; GTPase\_Ras.  
DR InterPro; IPR003578; GTPase\_Rho.  
DR InterPro; IPR002041; RAN.  
DR InterPro; IPR001806; Ras\_trnsfrmg.  
DR InterPro; IPR002078; Sig54\_interact.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR SMART; SM00176; RAN; 1.  
DR SMART; SM00173; RAS; 1.  
DR SMART; SM00174; RHO; 1.  
DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; 1.  
KW Hypothetical protein.  
FT NON\_TER  
SQ SEQUENCE 211 AA; 23761 MW; 3BFA7CF89D4DF54F CRC64;

Query Match 51.9%; Score 573; DB 4; Length 211;  
Best Local Similarity 62.8%; Pred. No. 9e-41;  
Matches 125; Conservative 15; Mismatches 39; Indels 20; Gaps 4;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKIRIQ 60  
Db |||||  
Qy 4 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKIRIQ 63  
Db |||||  
Qy 61 IWDTAGERYQTITKQYRRAGGIFLVYDIDISSERSYQIMKWSVDVDEYAPGQKILIG 120  
Db |||||  
Qy 64 IWDTAGERYQTITKQYRRAGGIFLVYDIDISSERSYQIMKWSVDVDE-----VG 113  
Db |||||

Qy 121 NKADEEQKQVREGGQOLAKKEYGMDFYETSACTNLNLIKESFTRLTFLVLAHQHKELEGL 180  
Db |||||  
Qy 114 DATS-----LPGGEGASPGKARRGPDGKANASRKLCLPQPMWKTSGTHQKASRRSLGI 168  
Db |||||  
Qy 181 R-MRANSELALABLEEBSG 198  
Db |||||  
Qy 169 RLMRSRN-----GRWEESKG 183  
Db |||||

## RESULT 3

Q7T3A4 PRELIMINARY; PRT; 200 AA.  
AC Q7T3A4;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN SEQUENCE FROM N.A.  
RP TISSUE=Kidney;  
RC MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;  
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;  
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;  
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Haieh F.;  
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;  
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;  
RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;  
RA Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.;  
RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;  
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;  
RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;  
RA Fahey J.; Helton E.; Kettelman M.; Madan A.; Rodrigues S.; Sanchez A.;  
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;  
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;  
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Butterfield Y.S.;  
RA Krzywinski M.I.; Skalka U.; Smallus D.E.; Schnerch A.; Schein J.E.;  
RA Jones S.J.; Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN SEQUENCE FROM N.A.  
RP TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC053195; AAH53195.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 200 AA; 22458 MW; FA177CEDAEF2820C CRC64;

Query Match 51.0%; Score 563.5; DB 13; Length 200;  
Best Local Similarity 58.0%; Pred. No. 5.4e-40;  
Matches 102; Conservative 44; Mismatches 29; Indels 1; Gaps 1;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKIRIQ 60  
Db |||||  
Qy 61 IWDTAGERYQTITKQYRRAGGIFLVYDIDISSERSYQIMKWSVDVDEYAPGQKILIG 120  
Db |||||  
Qy 61 IWDTAGERYQTITKQYRRAGGIFLVYDIDISSERSYQIMKWSVDVDEYAPGQKILIG 120  
Db |||||  
Qy 121 NKADEEQKQVREGGQOLAKKEYGMDFYETSACTNLNLIKESFTRLT-ELVLAHQH 175  
Db |||||  
Qy 121 NKCDIEAKRVSKETGEKLAKEHGFPEFSKSSINVEESFTSLARDILLKSNK 176  
Db |||||





RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL; AE003572; AAF50924.1; -;  
 DR EMBL; AB061899; BAA21744.1; -;  
 DR EMBL; AY060425; AAL25464.1; -;  
 DR HSP; P05713; 3RAB.  
 DR FlyBase; FBgn0015789; Rab10.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0005252; F:GTP binding; IEA.  
 DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
 DR GO; GO:0015031; P:protein transport; IEA.  
 DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . . ; IEA.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR002078; Sig54 interact.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMNG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; 1.  
 KW GTP-binding; Lipoprotein.  
 SQ SEQUENCE 204 AA; 23336 MW; 4E058761C6854920 CRC64;

Query Match 49.5%; Score 546.5; DB 5; Length 204;  
 Best Local Similarity 49.8%; Pred. No. 1.5e-38;  
 Matches 105; Conservative 47; Mismatches 50; Indels 9; Gaps 3;  
 QY 1 MAKOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSHISTIGVDFFKMTIEVDGKIVRI 59  
 DB 1 MAKTYDYLFRLLIGDSGVGKTCCLCRFTDNEPHSHISTIGVDFFKMTIEVDGKIVRI 60  
 QY 60 QIWDTAGQERYQTITTKOYRRAQIGIFLVYDISSRSYOHIMKWSVDVDEVAPEGVKIL 119  
 DB 61 QIWDTAGQERYQTITTKOYRRAQIGIFLVYDISSRSYOHIMKWSVDVDEVAPEGVKIL 120  
 QY 120 GNKADEQKRVGREGQOQLAKYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEG 179  
 DB 121 GNKCDMTDKRVNKGEGEATAREHGIFMETSAKSNINIERAFCELAAILD----KTS 176  
 QY 180 LRNRASNELALAELEEEEGEPGANSKTC 210  
 DB 177 RESAENQERVILDRNQEAP----GYSKCC 203

RESULT 6  
 Q8VCF6 PRELIMINARY; PRT; 207 AA.  
 AC Q8VCF6;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to mel transforming oncogene (Derived from cell line NK14)-RAB8 homolog (Cell line NK14 derived transforming oncogene).  
 DE oncoprotein.  
 GN MEL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Liver;  
 RA Strauberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Body, Cerebellum, and Retina;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 .RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
 DR EMBL; BC019990; AAH19990.1; -;  
 DR EMBL; AK076048; BAC36146.1; -;  
 DR EMBL; AK079306; BAC37603.1; -;  
 DR EMBL; AK080740; BAC38003.1; -;  
 DR HSP; P05713; 3RAB.  
 DR MGD; MGI:96960; Mel.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR InterPro; IPR003579; GTPase Rab.  
 DR InterPro; IPR001806; Ras trnsfrmg.  
 DR InterPro; IPR002078; Sig54 interact.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00071; ras; 1.  
 DR PRINTS; PR00449; RASTRNSFRMNG.  
 DR SMART; SM00175; RAB; 1.  
 DR TIGRFAMs; TIGR00231; small GTP; 1.  
 DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; 1.  
 KW GTP-binding; Lipoprotein.  
 SQ SEQUENCE 207 AA; 23668 MW; AC89DC85588FB8F8 CRC64;  
 Query Match 49.0%; Score 541; DB 11; Length 207;  
 Best Local Similarity 50.3%; Pred. No. 4.6e-38;  
 Matches 99; Conservative 49; Mismatches 49; Indels 0; Gaps 0;  
 QY 1 MAKOYDVLFRLLIGDSGVGKTCCLCRFTDNEPHSHISTIGVDFFKMTIEVDGKIVRI 60  
 DB 1 MAKTYDYLFRLLIGDSGVGKTCCLCRFTDNEPHSHISTIGVDFFKMTIEVDGKIVRI 60  
 QY 61 IWDTAGQERYQTITTKOYRRAQIGIFLVYDISSRSYOHIMKWSVDVDEVAPEGVKIL 120  
 DB 61 IWDTAGQERYQTITTKOYRRAQIGIFLVYDISSRSYOHIMKWSVDVDEVAPEGVKIL 120  
 QY 121 NKADEQKRVGREGQOQLAKYGMDFYETSACTNLNLIKESFTRLTLVLQAHKKELEG 180  
 DB 121 NKCDVNDKRVNKGEGEATAREHGIFMETSAKSNINVENAFRLARDIKAKMDKLEGN 180  
 QY 181 RMRASNELALAELEEE 197  
 DB 181 SPOGSSHGKVTVEQQK 197

RESULT 7  
 Q8KX35 PRELIMINARY; PRT; 203 AA.  
 AC Q8KX35;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE GTP-binding protein RAB13.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA Lau A.S.N., Cheng C.Y., Mruk D.D.;  
 RT "Rab13 participates in Sertoli cell tight junction assembly in the testis";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF525280; AA82588.1; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0005252; F:GTP binding; IEA.  
 DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
 DR GO; GO:0003930; F:RAS small monomeric GTPase activity; IEA.  
 DR GO; GO:0003931; F:Rho small monomeric GTPase activity; IEA.  
 DR GO; GO:0006886; P:intracellular protein transport; IEA.  
 DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . . ; IEA.  
 DR InterPro; IPR006688; ARF.

```
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR003577; GTPase Ras.
DR InterPro; IPR003578; GTPase rho.
DR InterPro; IPR002041; RAN.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54 interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00177; ARE; 1.
DR SMART; SM00175; RAB; 1.
DR SMART; SM00176; RAS; 1.
DR SMART; SM00173; RAS; 1.
DR SMART; SM00174; RHO; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding.
SQ SEQUENCE 203 AA; 22901 MW; B6687462C6EA95A4 CRC64;

Query Match 47.8%; Score 528.5; DB 11; Length 203;
Best Local Similarity 48.1%; Pred. No. 5.2e-37;
Matches 100; Conservative 53; Mismatches 46; Indels 9; Gaps 2;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCFTDNEFHSSHISTIGVDFKMTIEVDGKVRIQ 60
Db 1 MAKAYDHLFKLLIGDSGVGKTCLLIRFAEDNFNSTVISTIGIDFKIRTVEIEGKRIKQ 60

Qy 61 IWDTAGERYQTITKQYRRAGQIFLYVDISSERTSYOHIMKWSVDVEYAPGKVLIG 120
Db 61 VMDTAGQERFKTITATYRGANGILYVDITDEKSFENIQNMKSIIKENASAGVERLLG 120

Qy 121 NKADEOKRGVREGOQOLAKEYGMDFVETSACNLNLIKESFRTLTLVLQAHKKELEGL 180
Db 121 NKCDMEAKRVQERQAEHLAREHRIREFTSKSVNVDERAFSSLDIL----LKTGER 176

Qy 181 RMRASNELALAELEEEGKPEGPANSSK 208
Db 177 RSGNSKSPSSDLKVSDDK-----NSNK 199

RESULT 8
O18338 PRELIMINARY; PRT; 207 AA.
AC O18338
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE RAB8 protein (Lb44762p).
GN RAB8 OR CG8287.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yeaman M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Bereman B.P., Bhattacharya D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo J.B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
```

```
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON R; TISSUE=Head;
RX MEDLINE=97228579; PubMed=9074639;
RA Satoh A.K., Tokunaga F., Ozaki K.;
RT "Rab proteins of Drosophila melanogaster: novel members of the Rab-
RT protein family.";
RL FEBS Lett. 404:65-69(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; AE003515; AAF49101.1; -.
DR EMBL; AY069671; AAL39816.1; -.
DR EMBL; AY069671; AAL39816.1; -.
DR HSSP; P05713; 3RAB
DR FlyBase; FBgn0015796; Rab8
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. .; IEA.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54 interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 207 AA; 23691 MW; BE9DE812C77DCF09 CRC64;

Query Match 47.8%; Score 528.5; DB 5; Length 207;
Best Local Similarity 53.8%; Pred. No. 5.3e-37;
Matches 100; Conservative 41; Mismatches 40; Indels 5; Gaps 1;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCLLCFTDNEFHSSHISTIGVDFKMTIEVDGKVRIQ 60
Db 1 MAKTYDHLFKLLIGDSGVGKTCILRFSEDAFNFTTISTIGIDFKINTIEDLNKKIKQ 60

Qy 61 IWDTAGERYQTITKQYRRAGQIFLYVDISSERTSYOHIMKWSVDVEYAPGKVLIG 120
```

Db 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

RESULT 9  
ID Q9H0T3 PRELIMINARY; PRT; 200 AA.  
AC Q9H0T3;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN DKFZP564L1962.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=21154917; PubMed=11230166;  
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glaesl S.,  
RA Ansorge W., Boecker H., Bloeker H., Bauersachs S., Blum H.,  
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,  
RA Mewes H.W., Ottenwaelder B., Oberwalder B., Tampe J., Heubner D.,  
RA Wambutt R., Korn B., Klein M., Poustka A.,  
RA "Towards a Catalog of Human Genes and Proteins: Sequencing and  
RT Analysis of 500 Novel Complete Protein Coding Human CDNAs."  
RT Genome Res. 11:422-435(2001).  
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
RL  
DR EMBL; AL136650; CAB66595.1; -  
DR HSPSP; P05713; 3PAB.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
DR GO; GO:0015031; P:protein transport; IEA.  
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.  
DR InterPro; IPR003579; GTPase Rab.  
DR InterPro; IPR001806; Ras trnsfrmg.  
DR InterPro; IPR002078; Sig54 interact.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMs; TIGR00231; small GTP; 1.  
DR PROSITE; PS00675; SIGMA54 INTERACT 1; 1.  
KW Hypothetical protein; GTP-binding; Lipoprotein.  
SQ SEQUENCE 200 AA; 22469 MW; F701DB88E46E3EA CRC64;  
Query Match 47.6%; Score 526.5; DB 4; Length 200;  
Best Local Similarity 55.8%; Pred. No. 7.5e-37;  
Matches 96; Conservative 45; Mismatches 30; Indels 1; Gaps 1;  
Qy 1 MAKQYDVLFRLLIGSGVGKTCLLCRFTDNEPHSHSHSTIGVDYFKMTIEVDGKVR 59  
Db 1 MAKQYDVLFRLLIGSGVGKTCVLPFRSDAFNTTFTSTIGIDFKIKTVELQGGKIKL 60  
Qy 60 QIWDTAQGERFHTTTSYRGAMGIMLVYDITNCKSPENISKWLNRIDNEHANEDVERML 120  
Db 61 QIWDTAQGERFHTTTSYRGAMGIMLVYDITNCKSPENISKWLNRIDNEHANEDVERML 120  
Qy 120 GNKADESKQVGRGEOQOLAKYGMDFYETSACTNINIKESFRLTELVLQ 171  
Db 121 GNKCDMDKRVVPKGGQIAREHGIRFFETTSKANINIEKAFUTLAEDILR 172

## RESULT 10

ID Q9DD03 PRELIMINARY; PRT; 202 AA.  
AC Q9DD03;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE 0610007N03Rik protein (RIKEN CDNA 0610007N03 gene).  
GN 0610007N03RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Kidney;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barah G.,  
RA Blake J., Boffelli D., Bojunga N., Carrinci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,  
RA Hayaishizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.  
DR EMBL; AK002303; BAB22000.1; -  
DR EMBL; BC027214; AAH27214.1; -  
DR HSPSP; P05713; 3PAB.  
DR MGD; MGI:1915578; 0610007N03RIK.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0005525; F:GTP binding; IEA.  
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.  
DR GO; GO:0015031; P:protein transport; IEA.  
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.  
DR InterPro; IPR003579; GTPase Rab.  
DR InterPro; IPR001806; Ras trnsfrmg.  
DR InterPro; IPR002078; Sig54 interact.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00071; ras; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR SMART; SM00175; RAB; 1.  
DR TIGRFAMs; TIGR00231; small GTP; 1.  
DR PROSITE; PS00675; SIGMA54 INTERACT 1; 1.  
KW GTP-binding; Lipoprotein.  
SQ SEQUENCE 202 AA; 22770 MW; SDF599432E228AC0 CRC64;  
Query Match 47.6%; Score 526; DB 11; Length 202;  
Best Local Similarity 54.1%; Pred. No. 8.4e-37;  
Matches 92; Conservative 45; Mismatches 33; Indels 0; Gaps 0;  
Qy 1 MAKQYDVLFRLLIGSGVGKTCLLCRFTDNEPHSHSHSTIGVDYFKMTIEVDGKVR 60  
Db 1 MAKQYDVLFRLLIGSGVGKTCVLPFRSDAFNTTFTSTIGIDFKIKTVELQGGKIKL 60



```
Db 12 YDYLKLLGDSGVGKSCLLRFSDSGFTTFTTTIGDKIRTIELDKRIKLIQIWD 71
Qy 65 AGQERYOTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILGNKAD 124
Db 72 AGQERFTITAYYRGAGILLVYDVTDESFNIRNIRNIEQHASDENVKILVGNKAD 131
Qy 125 -BEQKRVGREGQQLAKYGMDFYETSACTNLIKESFTRLTLVLQA-----HRKELE 178
Db 132 MDESRAVPKSKQALADEYGMKFFETSAKTNLNVVEVFFSIADIKQRLADTDARAEPQ 191
Qy 179 GLRMRASNE 187
Db 192 TIKINQSDQ 200

RESULT 13
ID Q40218 PRELIMINARY; PRT; 214 AA.
AC Q40218;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Putative ethylene-responsive small GTP-binding protein.
GN P0565A07.136.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids I; Fabales; Fabaceae; Papilionoideae; Lotaea; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
binding proteins encoded by cDNAs from Lotus japonicus, and expression
of corresponding mRNAs in developing root nodules.";
RL Plant J. 11:237-250(1997).
CC -I- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z73947; CAA98175.1; -.
DR HSP; P05713; 3RAB.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005252; F:GTP binding; IEA.
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; Ras_1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 214 AA; 23627 MW; 5E1A5E83505E50D7 CRC64;

Query Match 46.9%; Score 518; DB 10; Length 214;
Best Local Similarity 48.6%; Pred. No. 4.3e-36;
Matches 103; Conservative 38; Mismatches 55; Indels 16; Gaps 3;

Qy 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSTIGVDFKMTIEVDGKVKRIQIWD 64
Db 12 YDYLKLLGDSGVGKSCLLRFSDSGFTTFTTTIGDKIRTIELDKRIKLIQIWD 71
Qy 65 AGQERYOTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILGNKAD 124
Db 72 AGQERFTITAYYRGAGILLVYDVTDESFNIRNIRNIEQHASDENVKILVGNKAD 131
Qy 125 -BEQKRVGREGQQLAKYGMDFYETSACTNLIKESFTRLTLVLQA-----HRKELE 178
Db 132 MDESRAVPKSKQALADEYGMKFFETSAKTNLNVVEVFFSIADIKQRLADTDARAEPQ 191
Qy 179 GLRMRASNE 187
Db 192 TIKINQSDQ 200

RESULT 14
ID Q7XHP7 PRELIMINARY; PRT; 215 AA.
AC Q7XHP7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Putative ethylene-responsive small GTP-binding protein.
GN P0565A07.136.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0565A07.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005194; BAC80082.1; -.
SQ SEQUENCE 215 AA; 23993 MW; F3DAD303C533C7EE CRC64;

Query Match 46.9%; Score 518; DB 10; Length 215;
Best Local Similarity 47.5%; Pred. No. 4.4e-36;
Matches 103; Conservative 42; Mismatches 48; Indels 24; Gaps 4;

Qy 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSTIGVDFKMTIEVDGKVKRIQIWD 64
Db 12 YDYLKLLGDSGVGKSCLLRFSDSGFTTFTTTIGDKIRTIELDKRIKLIQIWD 71
Qy 65 AGQERYOTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILGNKAD 124
Db 72 AGQERFTITAYYRGAGILLVYDVTDESFNIRNIRNIEQHASDENVKILVGNKAD 131
Qy 125 -BEQKRVGREGQQLAKYGMDFYETSACTNLIKESFTRLTLVLQAHRKELEGLMR 183
Db 132 MDESRAVPKSKQALADEYGMKFFETSACTNLNVEGVFFSIADIKQ----- 179
Qy 184 ASNELALAELEEE-----GKPEGPANSS---KTCWC 212
Db 180 ---RLAETDSKTDRTIKINKPEGDAEATTLQKSACC 213

RESULT 15
ID Q40215 PRELIMINARY; PRT; 216 AA.
AC Q40215;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Putative ethylene-responsive small GTP-binding protein.
GN P0565A07.136.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids I; Fabales; Fabaceae; Papilionoideae; Lotaea; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
binding proteins encoded by cDNAs from Lotus japonicus, and expression
```

```
Db 132 MDESRAVPKSKQALADEYGMKFFETSAKTNLNVVEVFFSIADIKQRLADTDHKAAPT 191
Qy 179 GLRMRASNELALAELEEECKPEGPANSSKTC 210
Db 192 TLKINQDSRAAGC-----AANKSSCC 213

RESULT 14
ID Q7XHP7 PRELIMINARY; PRT; 215 AA.
AC Q7XHP7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Putative ethylene-responsive small GTP-binding protein.
GN P0565A07.136.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0565A07.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005194; BAC80082.1; -.
SQ SEQUENCE 215 AA; 23993 MW; F3DAD303C533C7EE CRC64;

Query Match 46.9%; Score 518; DB 10; Length 215;
Best Local Similarity 47.5%; Pred. No. 4.4e-36;
Matches 103; Conservative 42; Mismatches 48; Indels 24; Gaps 4;

Qy 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSTIGVDFKMTIEVDGKVKRIQIWD 64
Db 12 YDYLKLLGDSGVGKSCLLRFSDSGFTTFTTTIGDKIRTIELDKRIKLIQIWD 71
Qy 65 AGQERYOTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILGNKAD 124
Db 72 AGQERFTITAYYRGAGILLVYDVTDESFNIRNIRNIEQHASDENVKILVGNKAD 131
Qy 125 -BEQKRVGREGQQLAKYGMDFYETSACTNLIKESFTRLTLVLQAHRKELEGLMR 183
Db 132 MDESRAVPKSKQALADEYGMKFFETSACTNLNVEGVFFSIADIKQ----- 179
Qy 184 ASNELALAELEEE-----GKPEGPANSS---KTCWC 212
Db 180 ---RLAETDSKTDRTIKINKPEGDAEATTLQKSACC 213

RESULT 15
ID Q40215 PRELIMINARY; PRT; 216 AA.
AC Q40215;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Putative ethylene-responsive small GTP-binding protein.
GN P0565A07.136.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids I; Fabales; Fabaceae; Papilionoideae; Lotaea; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
binding proteins encoded by cDNAs from Lotus japonicus, and expression
```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 21, 2004, 11:58:09 ; Search time 2486 Seconds  
(without alignments)  
2546.574 Million cell updates/sec

Title: US-09-817-198C-2  
Perfect score: 1105  
Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEKGKPGFANSKTCWC 212

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPto\_spool\_p/US09817198/runat\_15032004\_101746\_20571/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastcap -SUFFIX=rat -MINMATCH=0.1 -LOOPFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09817198 @CGN 1 1 4337 @runat\_15032004\_101746\_20571 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	996	90.1	1091	10	BF160330	BF160330 601768601
2	988	89.4	1004	13	BUS17848	BUS17848 AGENCOURT
3	953.5	86.3	932	10	BF535642	BF535642 602054039
4	922	83.4	616	12	BI648588	BI648588 603277781
5	915.5	82.9	759	13	BU343508	BU343508 603523088
6	914	82.7	992	10	BF101730	BF101730 601753464
7	910	82.4	902	10	BF178163	BF178163 601809083
8	885	80.1	699	12	BI557933	BI557933 603237549
9	884	80.0	515	29	AY420960	AY420960 Homo sapi
10	882	79.8	627	9	AL863254	AL863254 AL863254
11	880.5	79.7	773	14	CK030691	CK030691 AGENCOURT
12	876	79.3	515	29	AY420961	AY420961 Pan trogl
13	872	78.9	515	29	AY420962	AY420962 Mus muscu
14	858.5	77.7	796	10	BF966292	BF966292 602286692
15	846	76.6	623	14	CB578521	CB578521 AMGNNUC:N
16	843	76.3	882	13	BU841314	BU841314 AGENCOURT
17	838	75.8	1100	10	BF181167	BF181167 601805830
18	833.5	75.4	3151	11	CNSLT118T	EX248046 human ful
19	829	75.0	616	12	BG247902	BG247902 602359819
20	816	73.8	1121	10	BF182001	BF182001 601805936
21	815	73.8	481	13	BX283729	BX283729 EX283729
22	799	72.3	814	12	BG762967	BG762967 602735022
23	792	71.7	660	12	BG974411	BG974411 602844184
24	777	70.3	911	12	BI649317	BI649317 603278051
25	756	68.4	618	10	BF150485	BF150485 uy86D05.Y
26	733	66.3	1767	12	BM462147	BM462147 AGENCOURT
27	727	65.8	485	10	BE913096	BE913096 601668580
28	706	63.9	667	12	BI685808	BI685808 603309467
29	704	63.7	843	12	BG865609	BG865609 602783763
30	693.5	62.8	577	12	BG703353	BG703353 602683115
31	687	62.2	652	9	AL886867	AL886867 AL886867
32	686	62.1	465	10	BF565114	BF565114 UI-R-B01-
33	670	60.6	479	10	BF320679	BF320679 uz55d10.Y
34	666	60.3	543	10	BG019552	BG019552 daa21e04.
35	663	60.0	539	10	BG022618	BG022618 uy51f07.Y
36	658.5	59.6	693	10	BE287683	BE287683 601093322
37	653	59.1	468	13	BY276887	BY276887 BY276887
38	651	58.9	485	13	BY255176	BY255176 BY255176
39	651	58.9	823	12	BM017178	BM017178 603643780
40	650	58.8	479	13	BY256967	BY256967 BY256967
41	649	58.7	385	12	BM936454	BM936454 UI-M-BH2
42	645	58.4	487	13	BY247617	BY247617 BY247617
43	639	57.8	416	13	BY269783	BY269783 BY269783
44	639	57.8	468	13	BY260718	BY260718 BY260718
45	638	57.7	483	13	BY243237	BY243237 BY243237

ALIGNMENTS

RESULT 1  
BF160330  
LOCUS BF160330  
DEFINITION 601768601F1 NCI\_CGAP\_Lu29 Mus musculus cdna clone IMAGE:3987745 5',  
mRNA sequence.  
ACCESSION BF160330  
VERSION BF160330.1 GI:11040541  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1091)



Qy 41 ilecglyValaspPheLysMetLysThrIleGluValaspGlyIleLysValArgIleGln 60  
 |||||  
 Db 273 ATCGGTGTTGACATTAAAGATGAAGACTATCGATAGACGGCATCAAAAGTGAAGATACAG 332  
 |||||  
 Qy 61 IletTpsPthrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80  
 |||||  
 Db 333 ATTTGGGACACAGCAGGGCAGGAGGATACAGACTATCAAAAGCAGTACTATCGCGCA 392  
 |||||  
 Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100  
 |||||  
 Db 393 GCCCAGGGAATATTTTGTAGCTACGACATAGCAGTGGGCTCTCTATCAGCATATCATG 452  
 |||||  
 Qy 101 LysTpsValSerAspValaspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120  
 |||||  
 Db 453 AAGTGGGTGAGTACGTTGGTGTAGTCCAGAGGAGTCCAGAGAGTCTTAATTTGGG 512  
 |||||  
 Qy 121 AsnLysAlaaspGluGlnGlnLysArgGlnValGlyArgGlnGlnGlnLeuAla 140  
 |||||  
 Db 513 AATAAGGCTGATGAAGACGAAACGGCAGGTGGGGAGAGCAGGCGCAGCTGGCT 572  
 |||||  
 Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
 |||||  
 Db 573 AAGAGTACGCGATGGACTTCTCGAAACAGTCCCTGCACCAACCTTAATATTAAAGAG 632  
 |||||  
 Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180  
 |||||  
 Db 633 TCCTTCACTCGTCTGACGGAGTGGTGTGACGGCCACAGGAAGAGTGGATGGTCTC 692  
 |||||  
 Qy 181 ArgMetArgAlaSerAsnGluLeuAlaLeu-AlaGluLeuGluGlu-GluGluGly-Lys 199  
 |||||  
 Db 693 CGAACACGTGCCAGCAGCTCGCAGCTGGGCCAGCTGGAGAGGACCCAGGCCAAA 752  
 |||||  
 Qy 200 ProGluGlyPro 203  
 |||||  
 Db 753 CCTGAGGGGCC 764  
 |||||

RESULT 3  
 BF535642 602054039F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4193456 5',  
 LOCUS mRNA sequence.  
 DEFINITION

ACCESSION  
 VERSION BF535642.1 GI:11623010  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 JOURNAL NIH-MGC http://mgi.nci.nih.gov/  
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9525 row: m column: 09  
 High quality sequence stop: 685.

FEATURES  
 source Location/Qualifiers  
 1..932  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="4193456"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_SG2"

/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:  
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
 dt. Average insert size 1.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."

## ORIGIN

Alignment Scores: 4.57e-103 Length: 932  
 Pred. No.: 953.50 Matches: 187  
 Score: 94.63% Conservative: 7  
 Percent Similarity: 91.22% Mismatches: 9  
 Best Local Similarity: 86.29% Indels: 2  
 Query Match: 10 Gaps: 1  
 DB:  
 US-09-817-198c-2 (1-212) x BF535642 (1-932)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20  
 |||||  
 Db 145 ATGCGGAAACAGTACGATGTGCTTCCGGCTACTGCTGATCGGGGACTCCGGGGTTGGC 204  
 |||||  
 Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40  
 |||||  
 Db 205 AAGACATGCTGTGTGCTCCCTTCCCGCAACAGAGTTCCACTCTCGCATATCTCCACC 264  
 |||||  
 Qy 41 IleGlyValaspPheLysMetLysThrIleGluValaspGlyIleLysValArgIleGln 60  
 |||||  
 Db 265 ATCGGTGTTGACATTAAAGATGAAGACTATCGATGACGGCATCAAAAGTGAAGATACAG 324  
 |||||  
 Qy 61 IletTpsPthrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80  
 |||||  
 Db 325 ATTTGGGACACAGCAGGGCAGGAGGTACAGACTATCAAAAGCAGTACTATCGCGCA 384  
 |||||  
 Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100  
 |||||  
 Db 385 GCCCAGGGAATATTTTGTAGCTACGACATTAGCAGTGGCTCTCTATCAGCATATCATG 444  
 |||||  
 Qy 101 LysTpsValSerAspValaspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120  
 |||||  
 Db 445 AAGTGGGTGAGTACGTTGGTGTAGTCCAGAGGAGTCCAGAGAGTCCAGAGAGTCTTAATTTGGG 504  
 |||||  
 Qy 121 AsnLysAlaaspGluGlnGlnLysArgGlnValGlyArgGlnGlnGlnLeuAla 140  
 |||||  
 Db 505 AATAAGGCTGATGAAGACGAAACGGCAGGTGGGGAGAGCAGGCGCAGCTGGCT 564  
 |||||  
 Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
 |||||  
 Db 565 AAGAGTACGCGATGGACTTCTAGCAAAACAGTGGCTCGCAACACCTTAATATTCAAGAG 624  
 |||||  
 Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180  
 |||||  
 Db 625 TCCTTCACTCGTCTGACGGAGTGGTGTGACGGCCACAGGAAGAGTGGATGGGCTC 684  
 |||||  
 Qy 181 ArgMetArgAla-SerAsnGlnLeuAlaLeuAla---GluLeuGluGluGluGlyLy 199  
 |||||  
 Db 685 CGAACACGTGCCAGCAGCTCGCAGCTGGGCCAGCTGGAGAGGACCTGTGAGAGCGAGCGA 744  
 |||||  
 Qy 199 sProGluGlyPro 203  
 |||||  
 Db 745 ACCTGATGGGCCA 757  
 |||||

## RESULT 4

BI648588 616 bp mRNA linear EST 12-SEP-2001  
 LOCUS 60327781F1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5317917 5',  
 DEFINITION mRNA sequence.

ACCESSION  
 VERSION BI648588.1 GI:15562824  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM

REFERENCE  
 1 (bases 1 to 616)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

# **AUTHORS** **TITLE** **JOURNAL** **COMMENT**

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L14M1806 row: a column: 22  
High quality sequence start: 2  
High quality sequence stop: 614.

## **FEATURES**

source  
1..616  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129,C57BL/6J,FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5317917"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam3"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo df.  
Library constructed by Life Technologies. Investigators  
providing samples: Lohar Hennighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."

## **ORIGIN**

### **Alignment Scores:**

Pred. No.: 1.39e-99 Length: 616  
Score: 922.00 Matches: 180  
Percent Similarity: 98.92% Conservative: 3  
Best Local Similarity: 97.30% Mismatches: 2  
Query Match: 83.44% Indels: 1  
DB: 12 Gaps: 0

US-09-817-198c-2 (1-212) x B1648588 (1-616)

Qy 28 PheThrAspGluPheHisSerSerHisThrIleGlyValAspPheLysMet 47  
Db 25 TTCACCGACACAGATTCACCTCTCGCATCTCCACCATCGGTGTGACTTTAAGATG 84  
Qy 48 LysThrIleGluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGln 67  
Db 85 AAGACTATCGATGTAGACGGCATCANAAGTCAGATACAGATTGGGACACAGAGGCGAG 144  
Qy 68 GluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuVal 87  
Db 145 GAGAGGTACAGACTATCACAAAGCACTACTATCGGCGAGCCCA-GGAATATTTTATGTC 203  
Qy 88 TyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAsp 107  
Db 204 TACGACATTAGCAGTGGCGCTCTATCAGCATATCAGTAAGTGGGTCAGTGGAT 263  
Qy 108 GluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGln 127  
Db 264 GAGTACGCTCCAGAGAGGTCAGAGATCTTAATGGGATAGGCTGATGAGAGCAG 323  
Qy 128 LysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTyrGlyMetAspPhe 147  
Db 324 AAACGGCAGGTGGGAGAGCAGCGGCGAGCGTGGCTAAGGAGTACGGCATGCTTC 383  
Qy 148 TyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGlu 167  
Db 384 TACGAACAAGTGCCTGCACCAACTTAATTAAGAGTCCTTCACTCTGCTGACGAG 443  
Qy 168 LeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGlu 187

Db 444 CTGGTGTGTCAGGCCACAGAAAGAGCTGGATGTTCTCCGAACACGTGCCAGCAACGAG 503  
Qy 188 LeuAlaIeuAlaGluLeuGluGluGluGluGlyProGluGlyProAlaAsnSerSer 207  
Db 504 CTCGCATGTCGCCAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 563  
Qy 208 LysThrCysTrpCys 212  
Db 564 AAGACTGTCTGGTGC 578

## **RESULT 5**

BU343508

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BU343508 759 bp mRNA linear EST 28-NOV-2002  
603523088P1 CSFQCHN68 Gallus gallus cDNA clone CHEST465022 5', mRNA  
sequence.  
BU343508  
BU343508.1 GI:25851509  
EST.  
KEYWORDS  
SOURCE  
ORGANISM  
Gallus gallus (chicken)  
Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE  
1 (bases 1 to 759)  
AUTHORS  
Boardman, P.B., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
Rong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
TITLE  
A Comprehensive Collection of Chicken cDNAs  
JOURNAL  
Curr. Biol. 12 (22), 1965-1969 (2002)  
MEDLINE  
22335534  
PUBMED  
12445392  
COMMENT  
Contact: Simon Hubbard  
Department of Biomolecular Sciences  
University of Manchester Institute of Science and Technology  
(UMIST)  
PO Box 88, Manchester, M60 1QD, UK  
Tel: 01612008930  
Fax: 01612360409  
Email: Simon.Hubbard@umist.ac.uk.  
Location/Qualifiers  
1..759

## **FEATURES**

source

/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="Compton Line 151"  
/db\_xref="taxon:9031"  
/clone="CHEST465022"  
/sex="Female"  
/tissue\_type="cerebellum"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="CSFQCHN68"  
/note="Organ: brain; Vector: pBluescript II KS(+); Site\_1:  
EcoRI; Site\_2: NotI; this normalized library was  
constructed from 1 million independent clones. cDNA  
synthesis was initiated using an oligo(dT) primer, using  
methylated C in the first strand synthesis reaction.  
Following this first strand reaction, double-stranded cDNA  
was bluntended, ligated to NotI adapters, digested with  
EcoRI, size-selected, and cloned into the NotI and EcoRI  
compatible sites of a custom modified MCS of the  
pBluescript (KS+) vector. The library was normalized in 2  
rounds using conditions adapted from Soares et al., PNAS  
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
(1996) 791, except that a significantly longer  
reannealing hybridization was used."

## **ORIGIN**

### **Alignment Scores:**

Pred. No.: 1.16e-98 Length: 759  
Score: 915.50 Matches: 176  
Percent Similarity: 90.91% Conservative: 14  
Best Local Similarity: 84.21% Mismatches: 18  
Query Match: 82.85% Indels: 1



```

RESULT 7
BF178163          902 bp      mRNA      linear      EST 31-OCT-2000
LOCUS             601809083f1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4039701 5',
DEFINITION        mRNA sequence.
ACCESSION         BF178163
VERSION           BF178163.1 GI:11056305
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 902)
                  NIH-MGC http://mgc.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM9320 row: f column: 22
                  High quality sequence start: 2
                  High quality sequence stop: 603.
FEATURES          Location/Qualifiers
                   source          1..902
                   /organism="Mus musculus"
                   /mol_type="mRNA"
                   /strain="C57BL/6J"
                   /db_xref="taxon:10090"
                   /clone="IMAGE:4039701"
                   /tissue_type="tumor, gross tissue"
                   /dev_stage="7 months"
                   /lab_host="DH10B"
                   /clone_lib="NCI CGAP Mam5"
                   /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                   Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
                   Library constructed by Life Technologies. Investigators
                   providing samples: Lothar Hennighausen/Robin Humphreys,
                   NIH"
ORIGIN
Alignment Scores: 6.91e-98 Length: 902
Pred. No.: 910.00 Matches: 177
Percent Similarity: 97.28% Conservativeness: 2
Best Local Similarity: 96.20% Mismatches: 4
Query Match: 82.35% Indels: 1
DB: 10 Gaps: 0

US-09-817-198C-2 (1-212) x BF178163 (1-902)

Qy 30 AsphenGluPheHisSerSerHisIleSerThrIleGlyValAspPheLysMetLysThr 49
Db 14 GACAAACGAGTTCCACTCTCCGATATCTCCACCATCGGTGTGACTTTAAGATGAAGACT 73

Qy 50 IleGluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGlnGluArg 69
Db 74 ATCGAGTAGACGGCATCAAGTGTATATACAGATTGGGACACAGAGGGCAGAGAGG 133

Qy 70 TyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlnIlePheLeuValTyrAsp 89
Db 134 TACCAGACTATCAAAAGCAGTACTATCGGAGGCCAGGGAATATTTTGTCTACGAC 193

Qy 90 IleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyr 109
Db 194 ATTAGCAGTAGCGCTCTCTATCAGATATCATGAAGTGGGTGAGTGGTGGTGGTGGTGGT 253

Qy 110 AlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArg 129

```

```

Db 254 GCTCCAGAGAGGAGTCAGAGATCCTTAATTGGGAATAAGCTGATGAAGAGCAGAAACGG 313
Qy 130 GlnValGlyArgGlnGlnGlnGlnGlnLeuAlaLys-GluTyrGlyMetAspPheTyrG1 149
Db 314 CAGTGGGAGAGAGAGAGGGGAGAGCTGGCTAACGGAGTAGCGCATGGACTTCTACGA 373
Qy 149 uThrSerAlaCyThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuVa 169
Db 374 AACAGTGGCTGCACCAACCTTAATATAAAGAGTCCTTCACTCGTCTGACGGAGTGGT 433
Qy 169 lLeuGlnAlaHisArgLysGluLeuGlnGlyLeuArgMetArgAlaSerAsnGluLeuAl 189
Db 434 GCTGCAGGGCCACAGACAGAGCTGGTGTCTCCGAACACGTGCCAGCAGCAGCTCGC 493
Qy 189 aLeuAlaGluLeuGluGluGluGlyLysProGluGlyProGluGlyProAlaAsnSerLysTh 209
Db 494 ACTGGCGAGCTGGAGAGGAGGAGCAACCTGAGGGCCAGCAACTCTTCNAAGAC 553
Qy 209 rCysTrpCys 212
Db 554 CTGCTGGTGC 563

RESULT 8
BI557933
LOCUS             603237549f1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5290284 5',
DEFINITION        mRNA sequence.
ACCESSION         BI557933
VERSION           BI557933.1 GI:15445247
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 699)
                  NIH-MGC http://mgc.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM11734 row: b column: 13
                  High quality sequence stop: 698.
FEATURES          Location/Qualifiers
                   source          1..699
                   /organism="Mus musculus"
                   /mol_type="mRNA"
                   /strain="129, C57BL/6J, FVB/N"
                   /db_xref="taxon:10090"
                   /clone="IMAGE:5290284"
                   /tissue_type="tumor, gross tissue"
                   /dev_stage="10 months"
                   /lab_host="DH10B"
                   /clone_lib="NCI CGAP Mam3"
                   /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                   Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
                   Library constructed by Life Technologies. Investigators
                   providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                   Reference for transgenic model: Xu et al., Nature Genetics
                   22, 37-43 (1999)."
ORIGIN
Alignment Scores: 4.5e-95 Length: 699
Pred. No.: 885.00 Matches: 172
Score: 98.87% Conservativeness: 3
Percent Similarity:

```









```

142 Db -----
48 Qy LysThrIleGluValAspGlyIleLysValArgIleGlnIleTyrAspThrAlaGlyGln 67
160 Db AAGACTATCGAAGTAGACGGCATCAAGTGAAGATACAGATTGGGACACAGCGGCAG 219
68 Qy GluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuVal 87
220 Db GAGAGGTACACAGACTATCACAAAGCAGTACTATCGCGAGCCAGGGAATATTTTAGTC 279
88 Qy TyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTyrValSerAspValAsp 107
280 Db TACGACATTAGCAGTGAAGCGCTCTCATCAGCATATCATGAAGTGGTCAAGTGGAT 339
108 Qy GluTyrAlaProGluGluValGlnLysIleLeuIleGlyAsnLysAlaAspGluGluGln 127
340 Db GAGTACGCTCCAGAAAGAGTCCAGNAGATCTTAATTTGGGATATAGGCTGATGAAGAGCAG 399
128 Qy LysArgGlnValGlyArgGluGlnGlnGlnLysGluTyrGlyMetAspPhe 147
400 Db AAACGGCAGGTGGGAGAGAGCAGGGGCGACGCTGGCTAAGGAGTACCGCATGCACTTC 459
148 Qy TyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGlu 167
460 Db TACGAACAAGTGGCTGCACCAACCTTAATATTAAGAGTCTCTTCACTCGTCTGACGGAG 519
168 Qy LeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGlu 187
520 Db CTGGTGTCTGACAGGCCACAGAAAGAGCTGGATGTCTCCGAAACACGTCGACCAACAG 579
188 Qy LeuAlaLeuAlaGluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSer 207
580 Db CTGCACCTGGCGAGCTGGAGGAGGACGACGAGGACCAACCTGAGGGCCACGAACTCTTCA 639
208 Qy LysThrCysTyrTrpCys 212
640 Db AAGACCTGCTGGTGC 654

RESULT 12
AY420961 515 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM7398 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY420961
VERSION AY420961.1 GI:39776918
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
Location/Qualifiers
1..515
/organism="Pan troglodytes"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>515
/locus_tag="HCM7398"

gene
ORIGIN
Alignment Scores:
Pred. NO.: 3.31e-94 Length: 515
Score: 876.00 Matches: 169
Percent Similarity: 99.41% Conservative: 0
Best Local Similarity: 99.41% Mismatches: 1
Query Match: 79.28% Indels: 0
DB: 29 Gaps: 0

US-09-817-198c-2 (1-212) x AY420961 (1-515)
Qy 43 ValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIleTyr 62
Db 3 GTTGACTTTAAGATGAAGACCATAGAGTAGACGGCATCAAGTGGGATACAGATCTGG 62
63 AspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGln 82
Db 63 GACACTCGAGGGCAGGAGAGATACACAGCCATCACAAAGCAGTACTATCGCGGGCCAG 122
83 GlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTyr 102
Db 123 GGGATATTTTGGTNNNCGACATTAGCAGCGAGCGCTCTTACCAGCACATCATGAAGTGG 182
103 ValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLys 122
Db 183 GTCACTGACGTGGATGATGACGACCAAGAGGCGCTCCAGAGATCTTATTGGGAATAG 242
123 AlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGlu 142
Db 243 GCTGATGAGGAGCAGAAACGGCAGGTGGGAGAGAGAGAGGCGAGCGAGTGGCCAGAGAG 302
143 TyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPhe 162
Db 303 TATGGCATGGACTTCTATGAACAAAGTGCCTGCACCAACCTCAACATTAAGAGTCAATC 362
163 ThrArgLeuThrGluLeuValLeuAlaHisArgLysGluLeuGluGlyLeuArgMet 182
Db 363 ACGGCTCTGACAGAGCTGGTGTCTGACGGCCCATAGGAGGAGTGGAGGCGCTCCGGATG 422
183 ArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysProGluGly 202
Db 423 CGTGCCAGCATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGG 482
203 ProAlaAsnSerSerLysThrCysTyrCys 212
Db 483 CCAGCGAACTCTTCGAAACCTGCTGGTGC 512

RESULT 13
AY420962 515 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM7398 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY420962
VERSION AY420962.1 GI:39776919
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302

```

2 (bases 1 to 515)  
 Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence as made by sequencing genomic exons and ordering them  
 based on alignment.

FEATURES  
 source  
 1..515  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>515  
 /locus\_tag="HMCW7398"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,96e-94 Length: 515  
 Score: 872.00 Matches: 167  
 Percent Similarity: 99.41% Conservative: 2  
 Best Local Similarity: 98.24% Mismatches: 1  
 Query Match: 78.91% Indels: 0  
 DB: 29 Gaps: 0

US-09-817-198C-2 (1-212) x AY420962 (1-515)

Qy	43	ValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIleTrrp	62
Db	3	GTTGACTTTAAGATGAAGACTATCGAAGTAGAGCGCATCAAGTGAGATACAGATTGG	62
Qy	63	AspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAcgAlaGln	82
Db	63	GGACACAGCGGGCAGGAGAGGTHACGACTATCACAAAGCAGTACTATCGCGAGGCCAG	122
Qy	83	GlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrrp	102
Db	123	GGAAATATTTTAGTCTACGACATTAGCAGTGAGCGCTCTCATCAGCATATCATGAAAGTGG	182
Qy	103	ValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLys	122
Db	183	GTCAAGTACAGTGGATGAGTACGCTCCAGAAAGGAGTCCAGAAGATCTCTAATTGGGAATAAG	242
Qy	123	AlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGlu	142
Db	243	GCTGTGTAAGACAGAAACCGCAGTGGGAGAGAGCAGCGGGCAGCAGCTGGCTTAGGAG	302
Qy	143	TyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPhe	162
Db	303	TACGGCATGACTTCTACGAAACAAAGTGCTCGACCAACCTTAATATTAAAGAGTCTCTTC	362
Qy	163	ThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMet	182
Db	363	ACTCGTCTGACGGAGCTGGTGTGTCAGGCCCCACAGGAAGAGTGTGATGTCTCCGAACA	422
Qy	183	ArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysProGluGly	202
Db	423	CGTGCCACGACACAGCTCGCACCTGGCGAGCTGGAGGAGGACCAAGCAACCTTGAGGGC	482
Qy	203	ProAlaAsnSerSerLysThrCysTrrpCys	212
Db	483	CCAGCAACTCTTCAAAGACTGCTGCTGTC	512

RESULT 14  
 BF966292  
 LOCUS  
 DEFINITION  
 602286692F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:4375366 5',  
 mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 BF966292.1 GI:12333507  
 EST.

Db 309 AAGTGGGTCTAGTACGATGAGGTAGGAGATGCCACCTCAGTCGCCGGGTGTGGAGAG 369  
Qy 108 ----- 108  
Db 369 GGTGCTCACCGGGAAGCCAGCGAGGCCAGATGGGAGGCAATGCTTCAGGAAG 428  
Qy 109 -----TyrAlaProGluGlyValGlnLys11 117  
Db 429 CTTGCTTCCACAGCCCTGGATGAAGACCTCTGGTACGCCACAGAGGCGTCCAGAAAGAT 488  
Qy 117 eLeuileGlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyG1 137  
Db 489 CCTTATTGGGAATAAGGCTGATGAGGACAGAAACGCGAGGTGGGAAGAGAGCAAGGGCA 548  
Qy 137 nGlnLeuAlaLysGluTyrGlyMetAsp-PheTyrGluThrSerAlaCysThr-AsnLeu 156  
Db 549 GCAGCTGGGAGAGATGATGCGATGGACTTCTATGAACANGTCCCTGCACCAACCTC 608  
Qy 157 AsnIleLysGluSerPheThr-ArgLeuThrGluLeuValLeuAlaHisArgLys-- 175  
Db 609 AACATTAAAGAGTCATTTCACCGCGTCTGACAGAGCTGGTTCGACGGCCATAGGAAAGG 668  
Qy 176 GluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGlu 195  
Db 669 GACCTGGAAGCCCTCCGATCGCTGCCAGCAATGAGTTGGCAGCGGCAAGCTGGAGGAAC 728  
Qy 196 GluGluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTyr 211  
Db 729 GAGGAGGGCCAAACCGAGGGCCAGG-AACTTTGGAAAACCTGGGTGG 775

## RESULT 15

CB578521  
LOCUS CB578521 623 bp mRNA linear EST 03-APR-2003  
DEFINITION AMGNNUC:NRH7-00035-B1-A nrhy7 (10850) Rattus norvegicus cDNA clone  
nrhy7-00035-b1 5', mRNA sequence.

ACCESSION CB578521  
VERSION CB578521.1 GI:29522562

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 623)

Angen EST Program.

Angen Rat EST Program

Unpublished (2003)

CONTACT: Dan Fitzpatrick

Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00035 row: b column: 1.

Location/Qualifiers

source

1. 623

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="nrhy7-00035-b1"

/clone\_lib="nrhy7 (10850)"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat  
hypothalmus adult female Wistar rat avg. insert size 2.3  
kb fraction 6 and 7"

## ORIGIN

## Alignment Scores:

Pred. NO.:	1.73e-90	Length:	623
Score:	846.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	76.56%	Indels:	0
DB:	14	Gaps:	0

US-09-817-198C-2 (1-212) x CB578521 (1-623)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuileGlyAspSerGlyValGly 20  
Db 136 ATGCGAAACAGTACGATGTGCTATTTCGGCTACTGCTGATCGGGGACTCCGGGGTGGGC 195  
Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40  
Db 196 AAGACCTGCTGCTGCTATGCGCGCTTCACCGACACAGAGTTCCACTCTCGCATATCTCCACC 255  
Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60  
Db 256 ATCGTGTGACCTTTAAGATGAATAACCTCGAGTAGACGGCATCAAAGTGAGGATACAG 315  
Qy 61 IleTyrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80  
Db 316 ATTTGGGACACAGCAGCGGAGGAGGTACCAGACTATCACAAAACAGTACTACCGGCGA 375  
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100  
Db 376 GCCCAGGGAATATTTTTAGTCTACGACATTAGCAGCGAGCGCTCTTTATCAGCACATCATG 435  
Qy 101 LysTyrValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuileGly 120  
Db 436 AAGTGGGTCTAGTGCAGTGGATGATGCCCCAGAGGCGTCCAGAAAGATCTCTCATCGGG 495  
Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140  
Db 496 AATAAGGCTCAGCAAGAAACAGAAACGCGAGGTGGGAGAGAGCAAGGCGACGCTGGCT 555  
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160  
Db 556 AAGGAGTATGGCATGGACTTTTACGAAACAAGTGCCTGSCACCAACCTTAACATTAAAGAG 615  
Qy 161 SerPhe 162  
Db 616 TCCTTC 621

Search completed: March 21, 2004, 13:39:16

Job time : 2507 secs

**THIS PAGE BLANK (USPTO)**